HAINE

ELECTRON MICROSCOPE

SPON

THE ELECTRON MICROSCOPE

THE PRESENT STATE OF THE ART

THE ELECTRON MICROSCOPE

The Present State of the Art

by

M. E. HAINE

D.Sc.(Bristol), M.I.E.E., F.Inst.P.

Director of the Associated Electrical Industries Research Laboratory, Harlow

including a chapter on

Specimen Techniques and Applications

written in collaboration

with

V. E. COSSLETT

M.A.(Oxon.), M.Sc.(Lond.), Ph.D.(Bristol)

Cavendish Laboratory, Cambridge

1961

E. & F. N. SPON LIMITED

22 Henrietta Street, London, W.C.2

First published 1961
Reprinted 1962

Made and printed in Great Britain by
William Clowes and Sons, Limited, London and Beccles

ERRATA

p. ix line 9 should read '. . . Ruska and Knoll . . .' not '. . . von Borries and Ruska . . .'

p. 11 line 3 should read '. . . about 0·008.'

p. 16 Fig. 1.12 the abscissa scale should read 0 to 0·09

p. 30 Fig. 2.4 phase difference angle should read η_r

p. 32 Equation 2.2 should read

$$\psi' = \int_{-\pi}^{+\pi} \int_{0}^{\infty} \frac{\psi_r \eta_r r}{f\lambda} \exp\left[i\left(\eta_r + \frac{2\pi\Delta}{\lambda}\right)\right] \mathrm{d}r\, \mathrm{d}\xi$$

p. 35 bottom, second integration should read

$$\psi'' = \frac{\psi' r_0'}{r''} J_1\left(\frac{2\pi r_0' r''}{f\lambda}\right)$$

p. 71 line 5 '. . . assess the results of each . . .'

p. 74 line 3 '. . . where $e\phi_e$ is the loss in energy.'

p. 75 line 3 '. . . limited validity . . .'

p. 79 Table 4.1 'd' not 'd^2'

p. 82 Table 4.2 Z values should read $\frac{4}{16}$ $\frac{64}{64}$ not $\frac{4}{4}$ $\frac{64}{16}$

p. 120 Fig. 6.3 Add to caption 'The brightness figures refer to an electron energy of 50 keV'

p. 128 Fig. 6.8 $d = \cdot 05$ in. should be arrowed to second curve.

p. 136 Caption to Fig. 6.11 should read '. . . and two cathode shield geometries.'

p. 164 Fig. 8.4 should read 'Diagram of . . .'

p. 181 Equation 8.6 should read

$$L_3 \geqslant r_s \left(\frac{f}{S+D}\right)\left(\frac{4C_d}{x_0}\right)^{1/2}\left(1+\frac{S}{D}\right)$$

p. 181 line 23 delete 'V_r in'

p. 183 line 12 should read '. . . being refracted through . . .'

p. 183 para. 4 equations should read

$$r_2 = \left[\frac{xr_3 f_{21}{}^4}{L_2 C_s}\right]^{-1/3}$$

and

$$M_{\min} = \frac{r_s L_1}{f_1}\left[\frac{xr_3 f_{21}{}^4}{L_2 C_s}\right]^{-1/3}$$

p. 184 line 1 should read '. . . imposed that $r_3 \times$ . . .'

p. 184 Equation 8.10 should read

$$R_g = L_2\left[\frac{xL_1}{5r_3{}^2(L_1+L_2)f_2}\right]^{1/3}\left(\frac{f_2}{S+D}\right)^{-2/3}$$

p. 184 para. 2 equation should read

$$f_2 = \frac{M_1 M_3 L_2}{M_t}$$

p. 184 last equation should read

$$R_g = 1{\cdot}3\left[\frac{xf_1 M_t L_2{}^2}{r_3{}^2 M_3(L_1+L_2)}\right]^{1/3} \text{ (for } f_2/(S+D) = 0{\cdot}3)$$

p. 187 Fig. 8.11 should be inverted

p. 221 para. 5 lines 1–2 should read '. . . was first proposed and built by Knoll . . .'

p. 222 para. 2 lines 17–18 should read '. . . better resolution than 500 Å.'

p. 225 line 12 reference should read 'KNOLL, M. *Z. Tech. Phys.*, **16** (1935), 467.'

CONTENTS

PREFACE

THOUGH magnified images were formed with electrons by Ambrose Fleming before 1900, the real story of the electron microscope dates back to the 1920s. In 1925 Louis de Broglie postulated his hypothesis concerning the wave nature of material particles and in 1926 Busch published the first analysis of the lens effects of solenoidal fields. The conception of the electron microscope followed quickly after this. Its birthplace was Berlin. The first instrument was constructed by von Borries and Ruska who modified a demountable high-speed cathode ray oscillograph. This beginning has led to the continuous development of the instrument and its application, not only in the Berlin laboratories but in many others throughout the world.

The stringent requirements of the instrument and its application have extended the skill and ingenuity of designer and technician to the full. As it often happens improved understanding of the principles and limitations involved have led to the possibility of simpler and cheaper designs, as well as improved performance. The understanding of these principles and limitations is also of the greatest importance to the user of the instrument, and it has been the prime motive for the preparation of this text to explain these matters in the simplest yet fullest manner. Mathematics have been excluded, as far as possible, and replaced by data. References are given to relevant literature, but not comprehensively. The literature on the subject has become enormous and the author has chosen bibliographical references more in the terms of illustration than of expansion. Apologies may be due to many authors whose papers are not mentioned. Natural choice has been made of those papers prepared by the author's close associates, and illustrations are drawn largely from his colleagues' and his own work. Full acknowledgement should be made to the fact that the instrument has been at

least equally well developed in other groups, especially in Professor Ruska's laboratory in Germany, its birthplace; Philips Research Laboratory; Delft in Holland; and in the R.C.A. Laboratory in the U.S.A. This only begins the widespread catalogue of workers and laboratories who have contributed so much to the field. Special mention must be made of the late Professor von Borries who, with Ruska, built the first electron microscope; and of the late Dr. G. Liebmann who, as colleague to the author, was the originator of much basic work on electron lenses and a great source of inspiration to all.

December, 1959 M. E. HAINE

ACKNOWLEDGEMENTS

GRATEFUL acknowledgement is made for permission to reproduce the following:

Figures 1.13, 3.1, 3.4, 5.3, 8.13 (*Advances in Electronics and Electron Physics*, Academic Press); Figures 6.5, 6.7, 6.8, 6.9, 6.11, 9.1 (*British Journal of Applied Physics*); Figures 5.5, 5.6, 5.7, 5.8, 8.14 (*Journal of Scientific Instruments*); Figures 8.7, 8.8 (*Proceedings of the Physical Society*, Great Britain); Figures 3.5, 10.13 (*Journal of The Royal Microscopical Society*, Great Britain); Figures 10.7, 10.26 (*Proceedings of the Royal Society A*); Figure 10.12 (*Journal of the Institute of Metals*); Figures 10.16, 10.19 (*Journal of Ultrastructure Research*, Academic Press); Figure 10.17 (*Biochimica et Biophysica Acta*, Elsevier, Amsterdam); Figures 10.21, 10.22, 10.23 (*Journal of Molecular Biology*, Aberdeen University Press); Figure 10.25 (*Philosophical Magazine*).

INTRODUCTION

THE unaided eye is incapable of perceiving detail finer than about 0·1 mm in size. The microscope is, essentially, an instrument which produces enlarged images of fine detail so that these can be perceived. The wave nature of light sets a fundamental limit to the fineness of detail which the microscope can 'resolve'. This limit bears no relationship to the microscope magnification, which can be increased indefinitely by the addition of more lenses, or the use of stronger lenses. The finest detail which a microscope can resolve is termed its *resolving power*. The finest detail which is reproduced in a given image may, for various reasons, be larger than the resolving power of the instrument and, in order to differentiate between this and resolving power, it is termed the *image resolution*. Thus the resolving power, as the term suggests, expresses the capability, theoretical or practical, of the instrument to resolve, while the resolution describes the result actually obtained.

The basic component of the microscope is the lens. The optical properties of lenses are too well known to require description here. The basic optical system for the microscope comprises a series of lenses which, in turn, produce a series of images of progressively increased size. The brightness of the final image depends upon the intensity of radiation from the object entering the objective lens, and the total magnification (M). Since this brightness falls in proportion to $1/M^2$ it is important to collect as much of the illuminating radiation as possible to give a visibly bright image. To achieve this end the angular aperture of the objective lens, that is, the angle of the cone of radiation accepted by the lens, must be large. This can be defined in terms of the semi-angle (α) of the cone of light leaving each point in the object and entering the objective lens. It is well known that the aberrations of a lens become more

severe as this angle is increased, so that in order to collect more light more perfect lenses are required.

There is, however, a much more important reason for using large apertures. It has already been mentioned that the resolving power of the microscope is limited by the wave nature of the radiation. The effect of this is that an infinitely small radiating object will be imaged as a patch of radiation of radius (d), given by:

$$d = 0{\cdot}61\lambda/\sin\alpha$$

where λ is the wavelength of the illumination in the space between object and objective lens.

It is now seen that, for the best resolution, large values of α and short wavelength radiation are required. The wavelength of green light is about 5000 Ångström Units (Å); for ultraviolet light, the wavelength is about 2500 Å. Going further down the electro-magnetic spectrum, X-rays can have wavelengths down to 1 Å and less but, unfortunately, satisfactory lenses for X-rays have not yet been devised.

The possibility of deflecting a stream of charged particles, such as electrons in an electrostatic or magnetic field, just as light is deflected at the boundary of two media of differing refractive indices, suggests the possibility of focusing such streams. In 1926 H. Busch published the first theoretical analysis of such lenses, using fields of axial symmetry. The idea of an electron microscope followed naturally from such work. It might have been thought that such a microscope held the promise of a resolving power limited only by the infinitesimal size of the electron had not Louis de Broglie, in 1924, hypothesised a wave nature for such moving charged particles, a hypothesis which was amply verified by the classic experiments of Davison and Germer on electron diffraction. The wavelength of the particulate stream is given by the relation:

$$\lambda = h/mv$$

where h is Planck's constant, m the electron mass, and v its velocity. The velocity of the electrons is related to the

electrostatic potential (ϕ) through which it has been accelerated by equating the kinetic energy to potential energy:

$$\tfrac{1}{2}mv^2 = e\phi$$
$$v = \sqrt{(2e\phi/m)}$$

where e is the electron charge. Hence:

$$\lambda = h/\sqrt{2em\phi}.$$

Putting in values for h, e and m, and expressing λ in Å,

$$\lambda = \sqrt{(150/\phi)}\ \text{Å}$$

Thus, for $\phi = 100$ kV,

$$\lambda = 0{\cdot}04\ \text{Å}$$

This is 10^5 times shorter than the wavelength of light, and is considerably smaller than interatomic spacings in a solid (2–5 Å), hence it would appear to offer no particular limitation if wide aperture lenses could be made available.

In fact, the best lenses that have so far been devised have usable apertures limited, by aberrations, to semi-angle not exceeding about 1/100 rad., thus limiting the attainable resolving power for the electron microscope to a few Å, or about the atomic spacing in a solid.

It will be seen, from later chapters, that contrast considerations play an important part in deciding just how much use can be made of this resolution. In addition, the practical difficulties met in attempting to achieve the theoretically promised results will be described.

In order to make detail of order of size d visible to the eye, an enlarged image must be produced with the detail of size greater than the resolution of the eye (0·1–0·3 mm). When d is only a few Å in size, the magnification must be very large (approaching $\times 10^6$). How this magnification is achieved is of no material importance so long as the method does not impose any limitation beyond that set by the objective lens. Thus, part of the magnification can be obtained by photographic enlargement, often photographing an image at some intermediate magnification.

Such enlargement is, of course, limited by the 'grain' of the photographic plate.

Since the mean free path of electrons in air at atmospheric pressure is very small the whole microscope system must be highly evacuated. It is, of course, necessary to have the facility for readily changing the specimen and, usually, the photographic plates. To allow this it is necessary to pump the system continuously. To save time, air locks are often employed for the insertion of specimens and photographic plates.

The illuminating electron beam is normally produced by accelerating electrons, which are emitted from a hot tungsten cathode in an electron gun, with a negative d.c. accelerating voltage applied to the cathode. Because the focal length of a magnetic lens is dependent on the electron energy, as well as the lens excitation current, it is necessary to stabilise the accelerating voltage and lens current. For very high resolving powers, the degree of stability required is very stringent, a few parts in 10^5, or less, variation being allowable. For electrostatic lenses compensation for this effect takes place if the lens is operated from the same potential source as the gun.

Having thus briefly described the salient points of the electron microscope principles, the following chapters will deal, one by one, with the details of design and operation.

CHAPTER I

OPTICAL PROPERTIES OF ELECTRON MICROSCOPE LENSES

An electron, travelling with a velocity v in an electrostatic or magnetic field, is subject to an accelerating force. In the electrostatic field, this force is given simply by:

$$eE = e\frac{\mathrm{d}V}{\mathrm{d}s}$$

where e is the electronic charge, and E the field strength (equal to the space derivative of the potential $\mathrm{d}V/\mathrm{d}s$). The direction of the acceleration is that of the field.

In the magnetic field, the accelerating force is given by Hev, where H is the magnetic field strength. The direction of the force is now perpendicular to the direction of motion and that of the field. Clearly, knowing these simple relations and the field distribution in any given system, the trajectories of electrons through the fields can be calculated or computed. It is not intended to go into the methods of calculation or computation in this book; very adequate descriptions have appeared elsewhere (Liebmann, 1950). For an understanding of the electron microscope it is only necessary to appreciate the results of the computations, that is, the properties of the lenses themselves. These can be expressed with adequate accuracy in terms of simple formulae or families of graphs, as will be shown. It is, however, desirable to appreciate a little of the form of the electron trajectories in the lenses, since these differ in some respects from those in the more conventional light/optical systems. The most important difference lies in the continuous variation of effective refractive indices in electron lenses as compared with the discrete steps occurring at glass interfaces in compound optical lenses.

The simplest form of electrostatic lens is illustrated in Figure 1.1, and consists of two coaxial cylinders, or apertures, with a potential difference $(V_2 - V_1)$ between

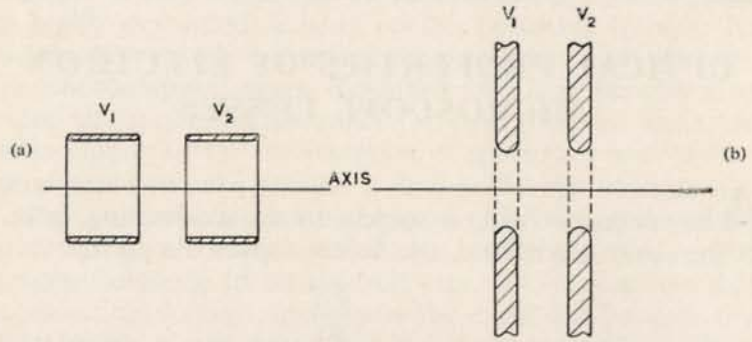


FIG. 1.1. *Diagrams illustrating the simple two cylinder* (a) *and two aperture* (b) *electrostatic lens.*

them. This type of lens is not convenient in the electron microscope because it requires different parts of the microscope to have different potentials. A more convenient arrangement (the Einzel lens) is that illustrated in Figure 1.2,

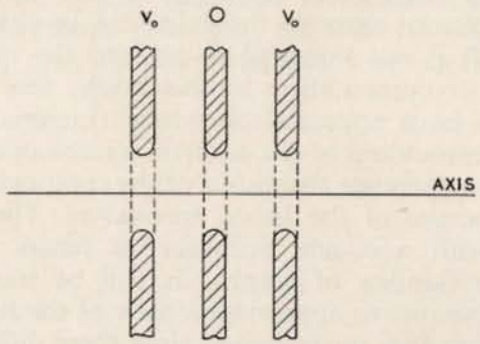


FIG. 1.2. *Diagram illustrating the three electrode Einzel lens.*

comprising three coaxial cylinders or apertures, the outer ones being connected to the microscope body (earth), and the middle one to the electron gun cathode. Figure 1.3

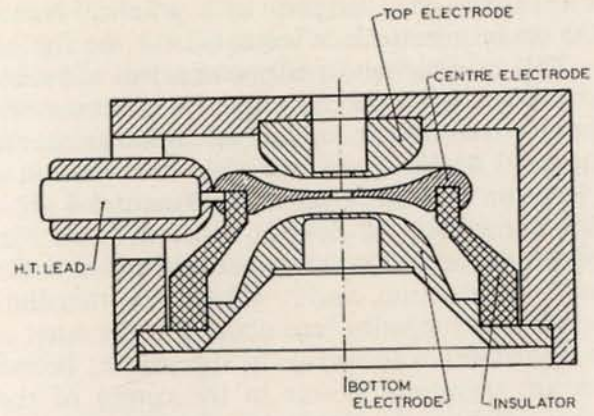


FIG. 1.3. *Cross-sectional diagram of typical Einzel lens.*

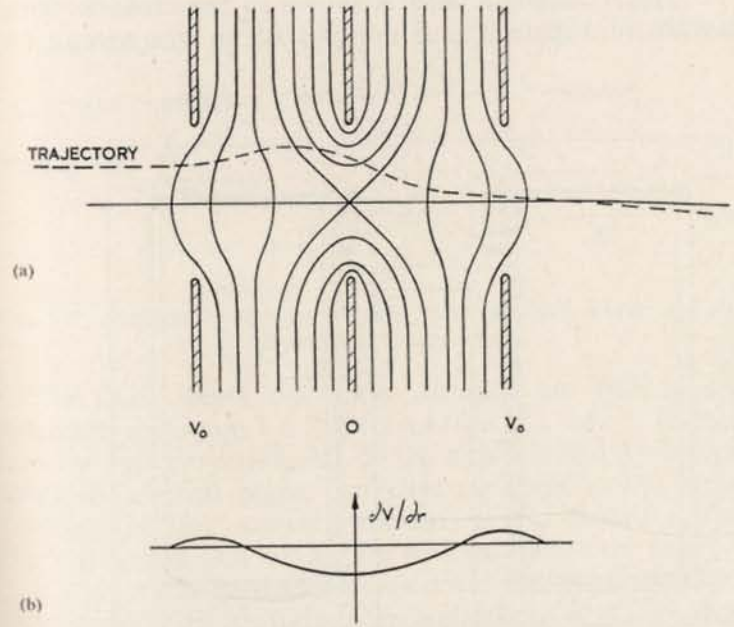


FIG. 1.4. (a) *Equal potential distribution and typical electron trajectory for three aperture Einzel lens.* (b) *Distribution of radial component of electron field in Einzel lens.*

shows a cross-sectional diagram of a practical lens of this type. The centre electrode is insulated for the full cathode voltage. This arrangement really comprises a double lens, the first half being a decelerating lens (because the field slows the electrons down), and the other half an accelerating lens. Figure 1.4 (a) shows the field distribution and a typical electron trajectory (dotted). Figure 1.4 (b) shows the radial component of field at some distance from the axis through the lens. It is, of course, the radial field which provides the lens action, and it will be seen that the radial field is outwards (negative lens action) at the outer ends of the lens and inwards (positive) in the centre; because the electrons are travelling slower in the centre of the lens, where the potential is closer to cathode potential, the inwards acting field has the greater effect so that the overall lens action is positive.

A typical magnetic lens is shown in the cross-sectional diagram of Figure 1.5. It consists of an iron circuit (A)

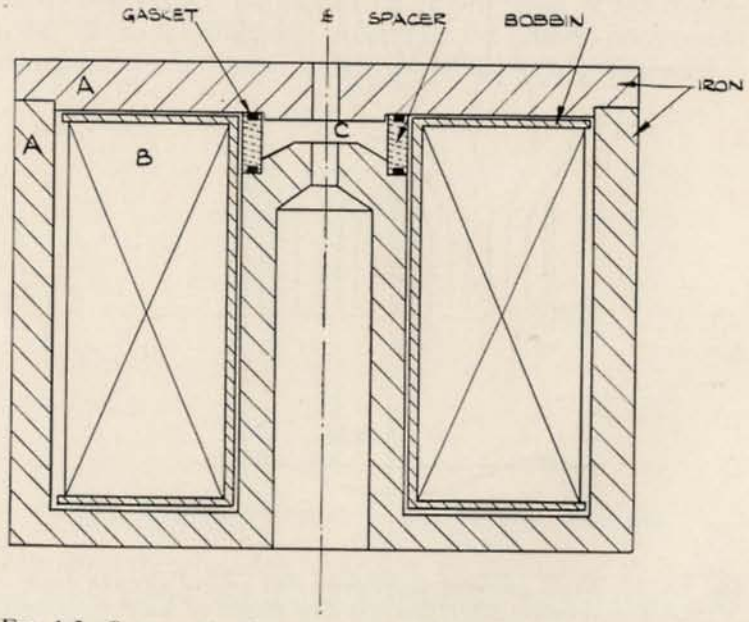


FIG. 1.5. *Cross-sectional diagram of typical magnetic electron lens.*

energised by a solenoidal coil (B). The iron circuit has a gap at (C) with parallel pole-faces bored with an axial hole which extends throughout the iron.

The important properties of electron microscope lenses are the focal lengths, the focal distances, the spherical and chromatic aberrations, the astigmatism arising from departures from symmetry (resulting from errors in manufacture) and, in the case of projector lenses, the aberration known as distortion. Other aberrations are of negligible magnitude.

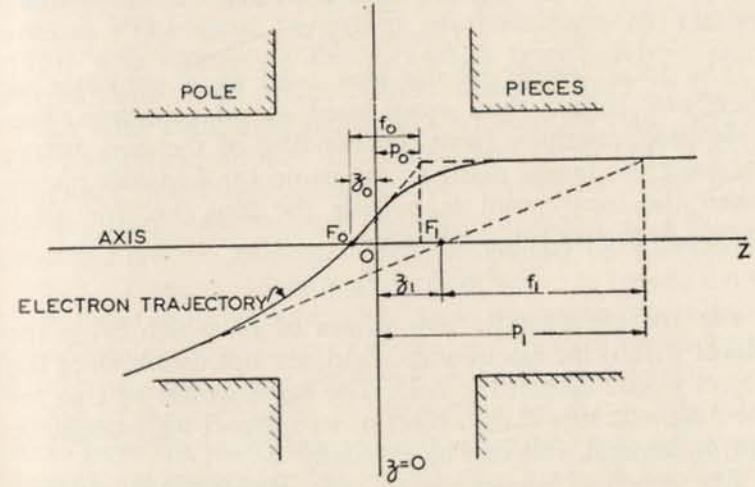


FIG. 1.6. *Diagram illustrating the definition of focal length and focal distance in electron lens.*

The focal length and focal distance are defined with reference to Figure 1.6 in which OZ is the axis of the lens, and the line perpendicular to the axis labelled $z=0$ represents the central plane. An electron trajectory is shown entering the lens, crossing the axis at F_0, distant z_0 from the lens centre, and leaving the lens parallel to the axis.

If the lens is used as an objective, the object would be positioned at F_0 for infinite magnification. It will be noted that this is very nearly the condition usually applying since stage magnifications of 100 or more are usual in the instrument. If a line is drawn tangential to the trajectory

at F_0, to cut the extended parallel outgoing ray at a distance p_0 from the lens centre, the plane through the intersection is the principle plane and the focal length for the objective lens is given by:

$$f_0 = p_0 + z_0.$$

When the lens is used as a projector the incoming ray, continued without deflection, would cross the axis at F_1, distant z_1 from the lens centre and, continued further, intersect the outgoing ray at distance p_1 from the lens centre. The effective focal length is then given by:

$$f_1 = p_1 - z_1.$$

The focal length of the lens used as a projector is, therefore, longer than when used as an objective. This difference, resulting from the bending of the rays before reaching the image plane in the projector lens, disappears when the focal point is outside the lens, i.e. for weak lenses. Similarly for weak lenses

$$f = f_1 = f_0 = z_1.$$

For the electrostatic lens values of f_0, which bring the object within the electrostatic field, are not usable since the object would distort the field. The same would be true for the magnetic lens if the object or its support were magnetic but, in general, this can be avoided.

The spherical aberration of any conventional electron microscope lens is positive, that is to say, the electrons traversing the lens far from the axis are refracted proportionately more than those close to the axis (paraxial). The magnitude of the spherical aberration is given in terms of the constant C_s, which relates the radial error δr of a ray making an angle α with the axis in the object plane when it arrives in the image plane by the relation:

$$\delta r = MC_s\alpha^3.$$

where M = the magnification.

Referring this error back to the object plane, the error becomes:

$$\delta r = C_s\alpha^3.$$

The combination of all such rays over an angular range 0–α_0 produces a caustic, the minimum radius of which is given by $\frac{1}{4}C_s\alpha^3$. The significance of this caustic is somewhat lost when the image formation is treated on a wave optical basis, as will appear in Chapter III. The significance of the constant C_s is, however, still unchanged and it is the variation of this constant with lens design that is required.

Focal Properties of Magnetic Lenses

Evaluations of the focal properties and aberrations for magnetic lenses have been carried out by a number of workers. The most important of these were by Glaser (1941) who calculated the properties from analytic field distributions fitted to computed distributions; by Lenz (1950) who computed fields and trajectories and, hence, lens properties; and by Liebmann and Grad (1951) who computed trajectories and, hence, properties from field distributions obtained with high accuracy ($\sim 1:10^4$) by the resistor network analogue method devised by Liebmann (1950 a). There is close agreement between the results from these three workers and, also, with the more exacting of experimental measurements carried out by various workers. All the data given here are based on Liebmann's results.

One of the difficulties in presenting results of this type arises from the many design variables which may influence them. A not inconsiderable contribution made by Liebmann has been to show how the relevant variables can be strictly limited, and the results given in greatly simplified form. This leads to the possibility of a much easier appreciation of the data, and a clearer approach to optimum design.

The magnetic lens properties for a given electron energy can be fully defined in terms of three parameters, the bore diameter (D) of the pole pieces (Figure 1.7), the pole spacing (S), and the magnetic field strength H_p which would occur between the poles if they were parallel and unperturbed by the bore holes, and thus given by:

$$H_p = \frac{4\pi NI}{10S}$$

where NI is the ampere-turn excitation applied to the gap.

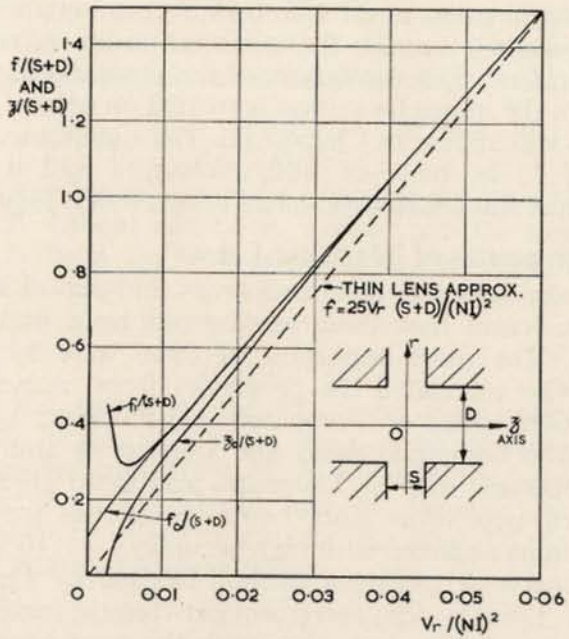


FIG. 1.7. *Curves showing the focal properties of magnetic lenses as functions of the excitation parameter* $V_r/(NI)^2$.

'Shaping' of the pole pieces has no significant effect, but increasing the excitation beyond some upper limit (where iron saturation becomes important) upsets the results as they will be given and, in all cases, produces disadvantageous effects in the lens properties. Liebmann (1953) showed that the upper useful limit of flux density could be specified by the useful 'rule of thumb' value where the permeability of the pole piece iron operating at a corresponding flux density falls below 50.

It will be seen that the lens can be defined by the value of one dimension such as S or D, the ratio of S/D and either H_p or NI. It happens that, if the values of f or z are plotted as a ratio $f/(S+D)$ or $z/(S+D)$ against an excitation parameter $V_r/(NI)^2$, where V_r is the relativistically-corrected electron energy, one line fully describes the properties (within an accuracy of 5%) for values of S/D

between 0·2 and 3·0. This range of the ratio S/D is adequate for any lens likely to be required in electron microscopy.

The relativistically-corrected electron voltage is related to the accelerating voltage V_0 by the relation:

$$V_r = V_0(1 + eV_0/2mc^2)$$

where c is the velocity of light, or:

$$V_r = V_0(1 + 0{\cdot}978 \times 10^{-6} V_0) \qquad (1.1)$$

Typical values are shown in Table 1.1.

TABLE 1.1

Relativistic correction of electron energy

V_0 (kV)	V_r (kV)
30	31
40	41·5
50	52·5
60	63·5
70	75·0
80	86·5
90	98
100	110
200	239
300	390

The 'universal' focal length and focal distance curves are shown in Figure 1.7. It will be seen that both curves are asymptotic to a straight line through the origin, and only depart significantly from it for values of $f/(S+D)$ or $z/(S+D)$ less than about unity. Lenses with this geometric ratio greater than unity are called weak lenses, and their focal lengths and focal distances are equal and given by:

$$f/(S+D) = z/(S+D) = 25V_r/(NI)^2 \qquad (1.2)$$

where V_r is measured in volts, and (NI) in ampere-turns.

It will be noted that the focal length and focal distance curves branch at the lower end into an upper and a lower part, giving the values for projector (subscript 1) and objective (subscript 0) lenses respectively.

It sometimes happens that circumstances require the use of an asymmetric lens in which the pole piece bores are unequal. In such a case, the universal curves still hold when D is replaced by the mean diameters of the two pole pieces.

Minimum Focal Length

It is often desirable to design a lens for the shortest possible focal length to give the maximum magnification. It will also be seen later that the strongest possible lens is required to minimise spherical aberration.

The minimum focal length for a given saturation flux density (H_{ps}) does not occur at the minimum of the $f/(S+D) - V_r/(NI)^2$ curve as might at first be expected. The conditions for the true minimum can be derived as follows:

Let the focal length curve of Figure 1.7 be represented by the function $\mathrm{F}[V_r/(NI)^2]$, then:

$$f/(S+D) = \mathrm{F}[V_r/(NI)^2].$$

The minimum value of S, as set by saturation, is given by:

$$S = 4\pi NI/10H_{ps}$$

Hence:

$$f_{min}H_{ps}/\sqrt{V_r}\left(1+\frac{1}{S/D}\right) = \frac{4\pi\ \mathrm{F}[V_r/(NI)^2]}{\sqrt{V_r/(NI)^2}}. \qquad (1.3)$$

Figure 1.8 shows the relation between the parameter on the left-hand side of this equation and the excitation

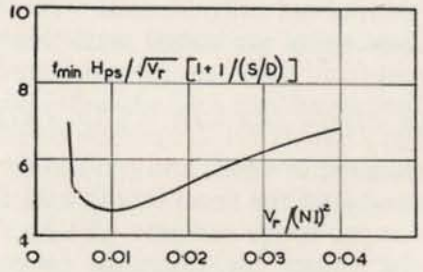


FIG. 1.8. *Relation giving minimum focal length as limited by iron saturation.*

parameter $V_r/(NI)^2$. It will be seen that the minimum focal length for a given voltage, field and ratio S/D occurs at a value of the excitation parameter of about 0·01.

For values of the ratio $S/D \gg 1$ the minimum focal length is given by:

$$f_{min} = 4{\cdot}5\sqrt{V_r}/H_{ps} \qquad (1.4)$$

e.g. for an electron energy of 100 keV and a saturation field of 18,000 gauss, $f_{min} = 0{\cdot}07$ cm.

It is seen that, for a given value of H_{ps} (determined by the properties of the iron used), and a given electron energy, there still remains an arbitrary choice for the ratio S/D. For minimum focal length, this ratio should be as large as possible, though little gain is obtained once S/D becomes large compared with unity. It usually transpires that practical difficulties limit the choice of the optimum ratio of S/D, since this calls for very small bore diameters, which are difficult to machine accurately, and often introduce other difficulties such as the insertion of the object.

Spherical Aberration of Magnetic Lenses

The value of C_s for the magnetic lens cannot be expressed quite as simply as that of the focal length, but a plot of C_s/f against $f/(S+D)$ shows only a small dependence on the ratio S/D. In Figure 1.9 values of C_s/f are plotted against $f/(S+D)$. The curves are accurate to within about $\pm 10\%$ for S/D between 0·2 and 2. It is seen that the spherical aberration increases very rapidly with focal length for a given lens size. In the electron microscope this rapid increase is largely offset by the fact that the instrumental resolving power is only dependent on the quarter power of the spherical aberration constant ($C_s^{1/4}$) as will appear in Chapter III.

For the thin lens condition the spherical aberration constant is given by the simple formula:

$$C_s/f = 5{\cdot}0[f/(S+D)]^2 \qquad (1.5)$$

$$= 3130V_r^2/(NI)^4. \qquad (1.6)$$

The minimum value of C_s/f is about 0·4 but, for practical reasons, including economy of excitation, a minimum value of about 0·7 is more realistic. This occurs at a value of $f_0/(S+D)=0{\cdot}3$, and $V_r/(NI)^2=0{\cdot}005$. For $V_r=100$ kV, an excitation of about 4500 ampere-turns is then

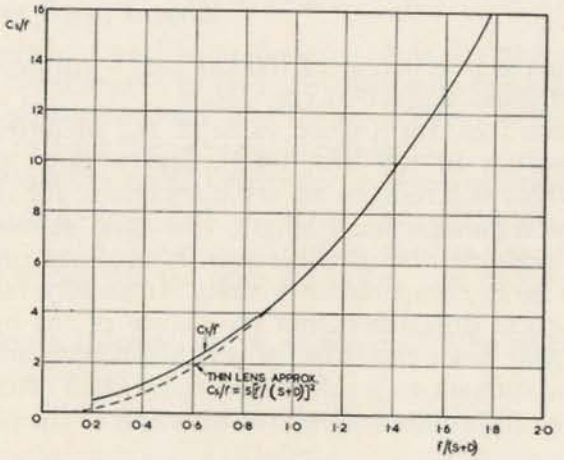


FIG. 1.9. *Variation of the ratio C_s/f with $f/(S+D)$.*

required. For a limiting flux density of 18,000 gauss, the spacing (S) would be 0·3 cm. For $S/D=1$, f now becomes 0·18 cm, and $C_s=0{\cdot}13$ cm. The choice of most suitable lens design will be discussed in more detail in Chapter III, but this example is quoted here for later comparison with the electrostatic lens.

Chromatic Defect

Since the focal length of the magnetic lens is dependent on both the electron energy and the lens excitation current, any variation in these parameters during operation of the microscope will result in the image focus being subject to a variation, thus resulting in a loss of resolution.

For a lens of given geometry, the focal length in the absence of iron saturation is given by:

$$f = KV_r/I^2$$

where K = a constant. By differentiation, we get:

$$\delta f = f(\delta V_r/V_r - 2\delta I/I). \qquad (1.7)$$

This equation relates the change in focal length (δf) to changes in electron energy or excitation current and therefore, if the maximum permissible change in focal length is specified, then the limiting permissible changes in voltage and current are determined.

To deal with the departure from linearity between f and V_r/I^2 in strong lenses, it is convenient to specify a new constant (C_c)—the chromatic constant—which, when replacing f in equation (1.7), makes the equation hold for all lenses. Thus:

$$\delta f = C_c(\delta V/V - 2\delta I/I). \qquad (1.7a)$$

The value of C_c has been computed by Liebmann, and is shown in ratio with f as a function of the excitation parameter in Figure 1.10. It will be seen that the ratio

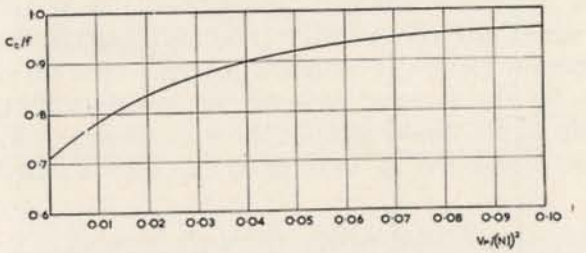


FIG. 1.10. *Variation of the ratio of chromatic constant (C_c) to focal length with the excitation parameter $V_r/(NI)^2$.*

C_c/f drops from unity for weak lenses to a value of about 0·7 for the strongest lenses. The limiting values of δf and the criterion for voltage and current stability will be further discussed in Chapter III while the influence of the chromatic defect on contrast will be referred to in Chapter IV. It is sufficient here to point out that the shorter the focal length the wider is the tolerance range for the voltage and current fluctuations.

Image Rotation

The magnetic lens has the peculiar property that the image is rotated through an angle round the axis with respect to the object. The rotation angle (ψ) in the 'thin lens' approximation is given by:

$$\psi \propto \int H_z \, dz$$

where H_z is axial magnetic field, and the integration is carried out from object to image.

This expression gives a proportionality between ψ and the ratio $NI/\sqrt{V_r}$. The curves of Figure 1.11 show that

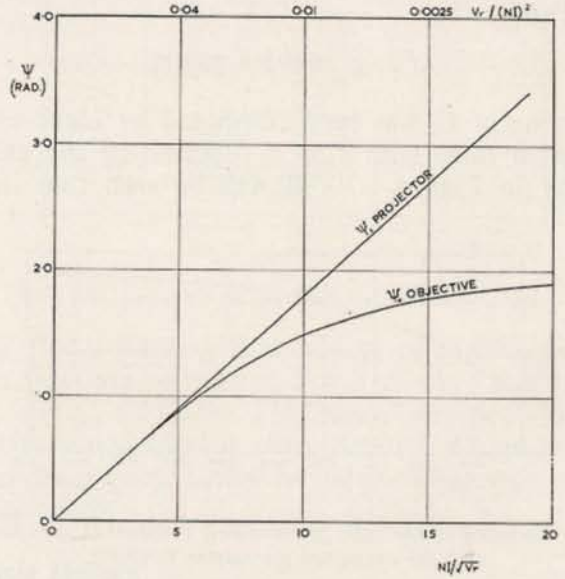


FIG. 1.11. *Image rotation angle in radians plotted as a function of* $NI/\sqrt{V_r}$ *and* $V_r/(NI)^2$ *(on top scale).*

this proportionality is closely followed, even for very strong lenses, particularly in the case of projector lenses. Objective lenses begin to depart from linearity at strengths rather less than that for minimum focal length. The curves shown in the figure are accurate to within a few per cent for $0{\cdot}2 < S/D < 2$.

The linear relationship is given by:

$$\psi = 0{\cdot}18NI/\sqrt{V_r} \text{ radians} \qquad (1.8)$$

Distortion

Distortion aberration is of particular importance in the final projector lens of the microscope, but can also be important in an intermediate projector when this is used under conditions of low magnification. In the magnetic lens, the image distortion may be resolved into two components, the first a radial change in magnification giving the well-known 'pin-cushion' or 'barrel' distortion, and the second a rotational distortion due to a change in the image rotation angle with radial distance from the axis. The former may be specified in terms of the fractional radial displacement of an image point from its undistorted position ($\delta\rho/\rho$), and the latter by the circumferential displacement, measured as a fraction of the radial position of the image point from the axis.

For a given geometry (S/D) and relative power $[f/(S+D)]$, it will be expected that the distortion will be dependent on some function of the ratio of the image point radius to the lens size, for example $F(\rho/MR)$, where M is the magnification and R the lens bore radius ($R = D/2$). The function can be expanded into a power series, the odd terms of which must, by conditions of symmetry, be zero. Hence:

$$\delta\rho/\rho = A(\rho/MR)^2 + B(\rho/MR)^4 + . \qquad (1.9)$$

In practice only the first term is of importance, and, for the constant A, is written either the radial distortion constant C_d or the spiral distortion constant C_{sp}, giving:

$$\delta\rho/\rho = C_d(\rho/MR)^2 \quad \text{(radial)} \qquad (1.10)$$
$$\delta\rho/\rho = C_{sp}(\rho/MR)^2 \quad \text{(spiral)} \qquad (1.11)$$

Liebmann has given values for C_d and C_{sp} but it will be shown, in Chapter VIII, that the quantities of direct interest in obtaining an optimum design are the product $\sqrt{C_d}(1+S/D)$, and $\sqrt{C_{sp}}(1+S/D)$, and these quantities are plotted as a function of the excitation parameter in Figure 1.12.

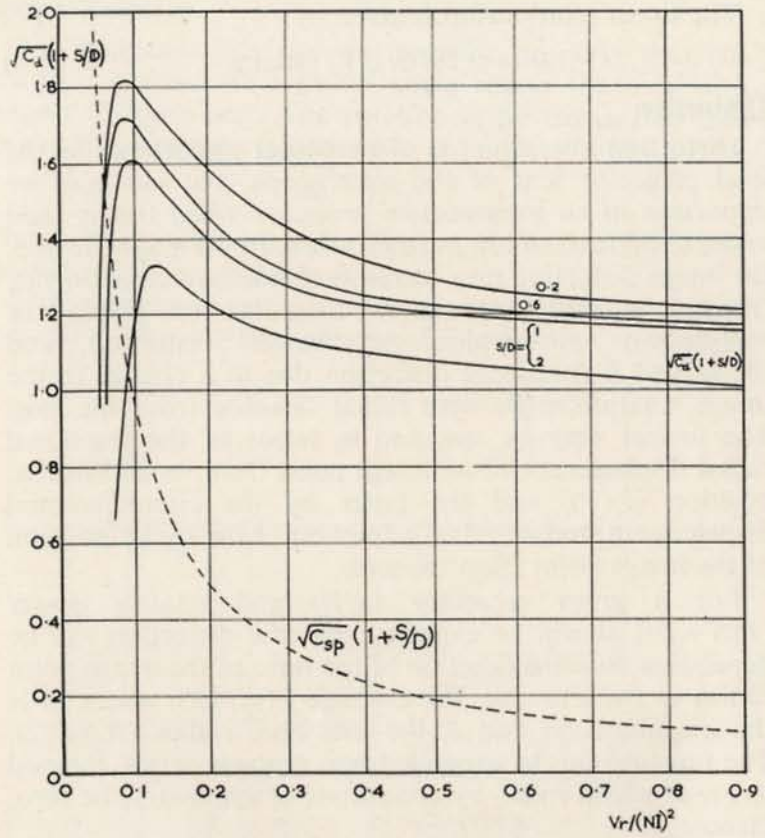


FIG. 1.12. *Variation of the distortion terms* $\sqrt{C_d}(1+S/D)$ *and* $\sqrt{C_{sp}}$ $(1+S/D)$ *with the excitation parameter* $V_r/(NI)^2$.

Astigmatism

It was pointed out by Hillier and Ramberg (1947) that mechanical or magnetic asymmetries in lens pole pieces or electrodes resulted in astigmatism in the image, and it has now been established that the mechanical tolerances on symmetry, which are required to reduce the astigmatism to negligible proportions for high resolution are finer than can be met by the most exacting manufacturing methods. Such defects, therefore, represent a factor which often limits the resolution of the electron microscope to a value

inferior to that as limited by spherical aberration and diffraction. Fortunately, it is possible to correct the astigmatism either by distorting the lens field in such a way as to correct the asymmetry, or by adding a weak cylindrical lens which gives an equal and opposite astigmatic effect. Before doing this, however, it is desirable to reduce the astigmatism as much as possible, so that the amount of correction to be applied is reduced to a minimum.

Bertein (1948), in an attempt to evaluate the relation between asymmetries and the resulting astigmatism, has expressed the asymmetries as a sum of Fourier components in the orientation angle, these components having periods π, $2\pi/3$, $\pi/2$, . . . , $2\pi/n$. He then calculates the effect of these asymmetry components in producing perturbations from axial symmetry in the field distribution, and hence the perturbations in the electron trajectories and the resulting aberration. A more complete calculation has been carried out for the magnetic lens by Sturrock (1951). Extending the methods of Bertein, he calculates fully the aberrations resulting from asymmetries in magnetic lenses. The asymmetries are resolved into four component types, illustrated in Figure 1.13. These are:

(a) misalignment—the lateral displacement of the axis of one pole piece with respect to the other;
(b) misorientation—a lack of parallelism between the axes of the two pole pieces;
(c) ellipticity—the departure from roundness of the pole piece bores;
(d) corrugation—the departure from flatness of the pole piece faces.

The conclusions from Sturrock's calculations show that the first two errors (a) and (b) result mainly in an effective bending of the lens axis, and hence a shift in the image, but only introduce a small degree of astigmatism and coma —one or two orders less in magnitude than that due to (c) and (d). On the other hand the last two asymmetries (c) and (d) result mainly in astigmatism and give no coma. Further, it is only the elliptical component of error (period π)

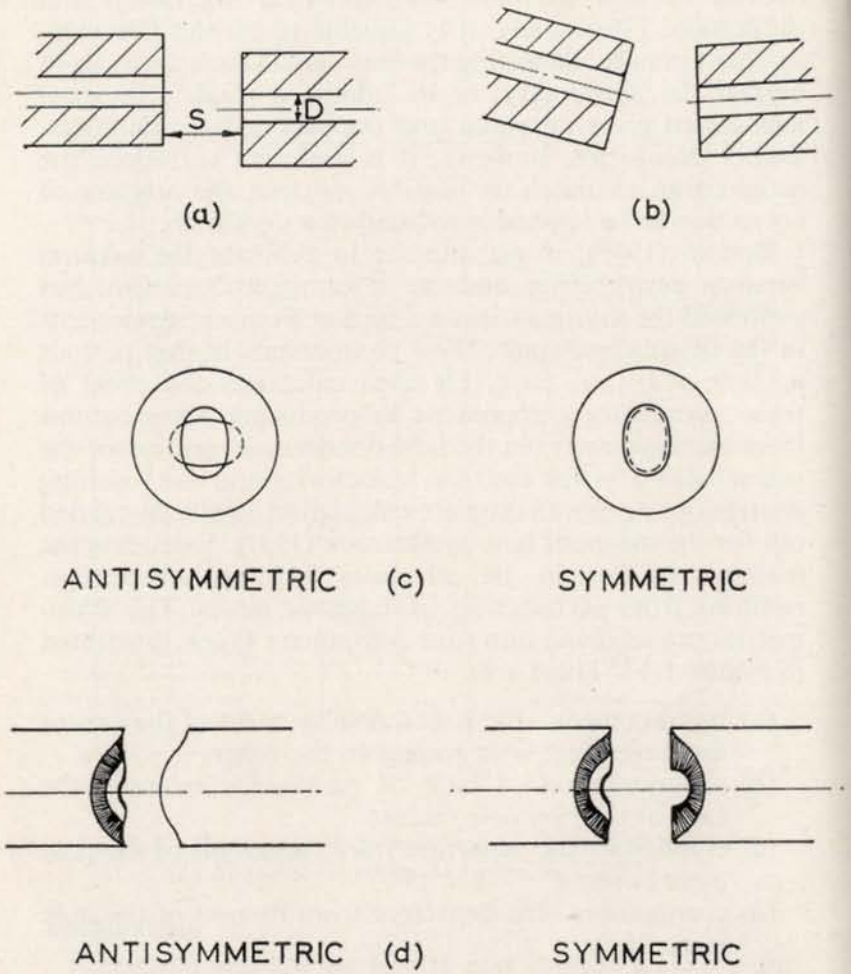


FIG. 1.13. *Illustrating the main component assymetries contributing to the production of astigmatism.* (a) *Misalignment;* (b) *Misorientation;* (c) *Ellipticity and* (d) *Corrugation.*

which is practically significant. The ellipticity of the top and bottom pole piece and the corresponding face corrugations may differ in amplitude and orientation. The total effect will be the vectorial sum of the effects due to each.

Calculations were carried out for two elliptic bores where the axes of cylindrical symmetry were parallel and also perpendicular and, in addition, for elliptic corrugations on the pole piece bores with their axes parallel and perpendicular.

Extensive computation on the basis of Sturrock's results were carried out by Archard (1953). From his results it is possible to select a maximum tolerance on the ellipticity and corrugation which must not be exceeded if the axial astigmatism (the axial distance between tangential and sagittal focii) is not to exceed a value z_a. The elliptic and corrugation errors can be expressed in terms of the peak to peak error (δ) in symmetry. The results which will be given refer specifically to ellipticity, but for all practical purposes they are accurate enough for application to corrugation errors also. It will be appreciated that great accuracy is not significant in practice since the error is an averaged one, in that it may vary along the bore with distance from the pole face, and with radius for the face corrugation. Also, the error may be different for the two pole pieces and the net effect will depend on the relative orientation of the error between one pole piece and the other. Furthermore, the requirement will be for the error to be negligible rather than that its value shall be precisely known. For these reasons an approximate assessment of its effect is sufficient. Archard's results show that the variation of the ratio z_a/δ with the excitation parameter $V_r/(NI)^2$ is linear and has a value given by:

$$z_a/\delta = 100(2+3S/D)[V_r/(NI)^2]. \qquad (1.12)$$

This expression is applicable even for strong lenses. For typical objective lens z_a/δ falls in the range 2–3. It will be shown, in Chapter III, that astigmatism requires a tolerance on symmetry often as low as a few tenths of a micron. Even in weak lenses, where only a very limited resolving power may be required, astigmatism can often be serious because of the higher ratio of z_a/δ under weak lens condition, as will appear in later discussions on condenser lenses.

Electrostatic Lenses

Data for electrostatic lenses are, unfortunately, less complete than for magnetic lenses. However, with some recent additions, there is enough data to make a close design evaluation for electron microscope purposes. The data must be presented in rather different form to that used for magnetic lenses, since the operation and the properties differ. In the first place, the electrostatic lens is excited with a voltage and, for all practical purposes, this voltage must be near cathode potential as the lens power falls off very rapidly with an applied potential below that of the cathode. It is, in any case, of advantage to tie the lens potential to that of the cathode as, by so doing, the focal length becomes closely independent of variations in the cathode potential, and hence the necessity for a highly stable voltage supply is eliminated. This fact represents a very important advantage of the electrostatic 'Einzel' lens which, as for all electrostatic lenses, also produces no image rotation.

The properties of the Einzel lens can be described in relation to the dimensional parameters described in Figure 1.14. S is the spacing between the outer (usually earthed) electrodes and the inner electrode, which is at cathode potential, D is the bore diameter of the central electrode, and T its thickness. The diameters of the bores of the holes in the outer electrodes are of little significance, provided they do not significantly exceed S. In practice the lens electrodes may be shaped somewhat to reduce surface electric stresses and voltage breakdown, but the effect of such shaping can usually be allowed for, in using the data, by choosing a reasonable equivalent ideal shape.

The variation of focal length and spherical aberration are shown in Figures 1.14 and 1.15 respectively. The data are due to Archard (1956). The dotted parts of the curves indicate that the focal plane is inside the lens field and in this region, therefore, the lenses cannot be used as objectives. It will be seen that an optimum ratio of T/D gives the minimum focal length and least spherical aberration.

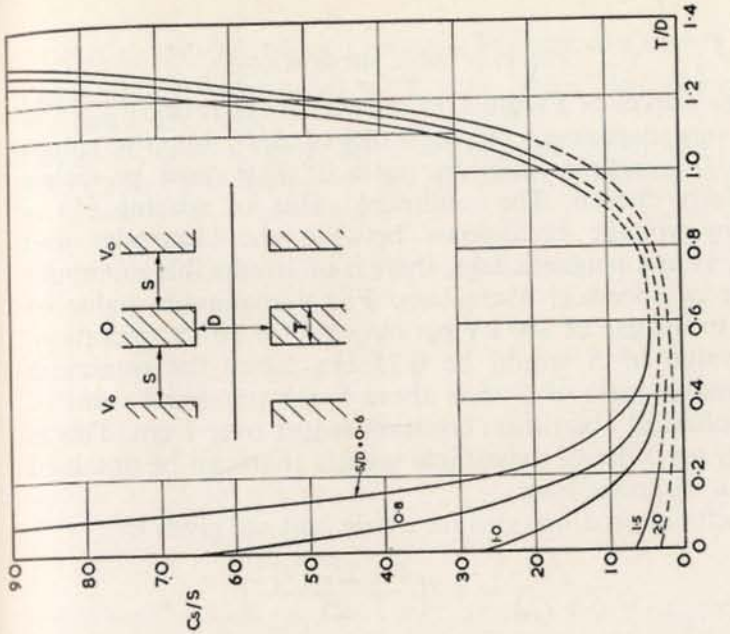


FIG. 1.15. *Spherical aberration of the three electrode Einzel lens.*

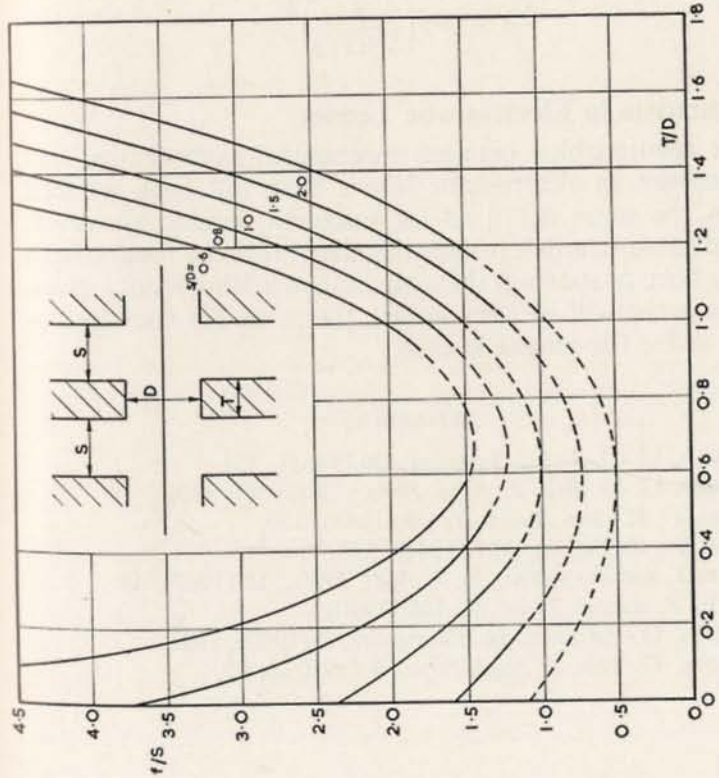


FIG. 1.14. *Focal properties of the three electrode Einzel lens as a function of its geometry.*

The curves of Figure 1.15 show that the ratio of C_s/S is not very dependent on the ratio of S/D, but for values of S/D less than unity the ratio of T/D must be rather carefully chosen. The minimum value of spacing (S) is set by voltage breakdown between the electrodes and thus, as in a magnetic lens, there is an irreducible minimum value of spherical aberration. For a maximum value of field in the gap of 200 kV per cm, and for 50 kV operation, the value of S would be 0·25 cm. Since the minimum obtainable ratio of C_s/S is about 5, the minimum value of the spherical aberration constant is just over 1 cm. This is nearly an order of magnitude greater than can be obtained with a magnetic lens.

Useful approximations for a thin lens are given by:

$$\frac{f}{S+T/2} = 2{\cdot}7\left(\frac{D+2S+T/2}{25}\right)^2$$

$$\frac{C_s}{f} = 3{\cdot}5\left(\frac{f}{S+T/2}\right)^2.$$

Astigmatism in Electrostatic Lenses

The relationships between mechanical asymmetries and astigmatism in electrostatic lenses have not been worked out in the same detail as for magnetic lenses. However, for the same strength of lens (i.e. same ratio of focal length to lens bore or spacing) the astigmatism arising from a given elliptic error will be very nearly the same for the electrostatic as for the magnetic case.

REFERENCES

ARCHARD, G. D. *J. Sci. Instrum.*, **30** (1953), 352.
ARCHARD, G. D. *Brit. J. Appl. Phys.*, 77 (1956), 330.
BERTEIN, F. R. *Ann Radioelec.*, **3** (1948), 379.
GLASER, W. *Z. Physik*, **117** (1941), 285.
HILLIER, J. and RAMBERG, E. *J. Appl. Phys.*, **18** (1947), 48.
LENZ, F. *Z. angew. Phys.*, p. 448 (1950).
LIEBMANN, G. *Advances in Electronics*, **2** (1950), 102.
LIEBMANN, G. *Brit. J. Appl. Phys.*, **1** (1950 a), 92.

LIEBMANN, G. and GRAD, M. E. *Proc. Phys. Soc.* (London), **54B** (1951), 956.
LIEBMANN, G. *Proc. Phys. Soc.* (London), **66B** (1953), 448.
STURROCK, P. A. *Trans. Roy. Soc.* (London), **A.243** (1951), 387.

CHAPTER II

ELECTRON WAVES, COHERENCE AND WAVE PROPAGATION

ALTHOUGH much of the theory of operation of the microscope can be explained in terms of geometrical optics, it is necessary to invoke the more accurate and comprehensive wave theory if a proper evaluation of the resolving power and contrast phenomena are to be attempted. For a proper understanding of the various phenomena which are met in using the instrument, it is desirable to have some understanding of the wave theory. In this chapter an attempt will be made to present the basic wave concept in the simplest possible form, making approximations wherever justifiable for the sake of simplicity.

The necessity for invoking the wave theory to explain the mechanics of particles is not at all obvious. Any attempt to justify it logically leads one immediately into philosophical argument far removed from the science of electron microscopy. The justification that must be accepted is largely experimental: a large number of experimental observations which cannot be explained by the simple theories of Newtonian mechanics are consistent with the wave theory. The dynamical theory often gives a solution which is apparently correct, but more precise experimental measurement shows this solution to be approximate; and, further, shows the wave theory to give a more accurate solution. The wave theory should be regarded as a more accurate theory rather than a quite different one. Thus, the first approximation in the result calculated by wave theory is often the solution given by the dynamic theory.

The Wave and Wave Propagation

In the wave theory, the propagation of a particle or stream of particles is explained by a system of waves, varying sinusoidally with time and propagated with constant velocity. Whether these waves have, in fact, any actual existence or not is irrelevant. The fact remains that the *assumption* of their existence, the calculation of their propagation, and the interpretation of the resulting distribution gives an answer which, up to the present, has always been consistent with observation; more consistent than the answer derived by dynamical calculation.

To introduce the concept of the wave, imagine a parallel stream of mono-kinetic electrons conveying an energy $\mathscr{E}$ per unit area per unit time, where $\mathscr{E}$ may vary across the beam. This situation is illustrated in Figure 2.1 by the

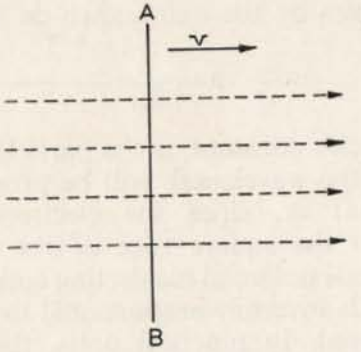


FIG. 2.1. *A plane wavefront, representing a parallel stream of particles.*

dotted parallel lines indicating the direction of flow (in optics ray paths). It will also be stated that this beam is coherent, though no significance will be applied to this term until later in the chapter.

No attempt will be made to derive the equivalent wave, but rather the result will be described. We may imagine the wave as a three-dimensional equivalent of the water wave. It will have at each point in space an amplitude varying sinusoidally with time, with a frequency $\omega/2\pi$; it will travel with a constant velocity v in free space, and

will therefore have a wavelength equal to the distance between one crest and the next. The phase of the oscillating plane wave will be constant across any normal cross-section of the beam, such as AB. Such a plane of constant phase is called a wavefront. A wavefront need not be a plane; for example, in a beam converging towards or diverging from a point, it will be spherical and it can have any other shape.

The most important property of the wave, for the present purpose, is the amplitude ψ which is everywhere proportional to the *square root* of the energy flow per unit area in the particle stream. The frequency and velocity are not of importance in normal optical calculations, except that the velocity is proportional to the reciprocal of the refractive index of the space traversed. The wavelength λ given by the ratio of the propagation velocity and the frequency is given by the well-known de Broglie relation:

$$\lambda = \frac{h}{mv}$$

where h is Planck's constant, m the particle mass, and v its velocity. Thus the wavelength will be proportional to the refractive index; or, since the electron velocities are proportional to the square root of the space potential relative to cathode potential (neglecting emission velocities), the wavelength is inversely proportional to the square root of space potential. In practical units, the wavelength is given by:

$$\lambda = \sqrt{\frac{150}{\Phi_0}}\ \text{Å}$$

where Φ_0 is the space potential (relative to zero electron energy) in volts.

The optical problem is: knowing the amplitude and phase distribution in one plane, or across one surface, to calculate the distribution across another surface some distance away—in other words, to calculate the propagation of the wave. The method depends upon the basic propagation principle in which each elemental area (δA) of the wave

may be considered as radiating a spherical 'wavelet' of amplitude $(\psi \delta A \cos \alpha)/s\lambda$, where ψ is the local wave amplitude, s the distance from the elemental area, λ the wavelength, and $\cos \alpha$ an obliquity factor, which, for the problems arising in the electron microscope, can almost always be put equal to unity. The solution of the propagation problem is reached by integrating the effects of all such elemental areas in the wave at every point on the distant surface under consideration, *taking into account the relative phases of the received wavelets*. The process is best illustrated by an example:

Suppose the amplitude ψ_0 and phase angle distribution η_0 are known at all points in a plane OR (Figure 2.2), in

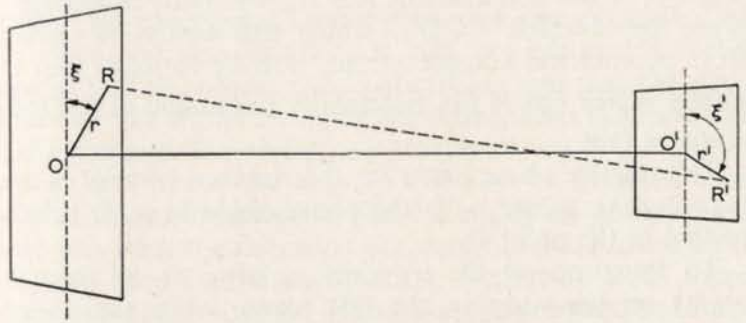


FIG. 2.2. *Diagram illustrating the derivation of path difference in a wave propagation.*

which a point is defined by the radial co-ordinate r, with zero at O, and orientation ξ, and it is desired to know the amplitude distribution in a parallel plane O'R'. The co-ordinates in the second place are taken as $(r'\xi')$, with zero at O', and the line OO' $(=s)$ is perpendicular to both planes.

Consider a point R' $(r'\xi')$, in the second plane: the amplitude of the wavelet emitted from an elemental area $r\,\mathrm{d}r\,\mathrm{d}\xi$ at co-ordinate position R, (r, ξ) at point R' will be given by:

$$\delta\psi = \frac{\psi_0}{\lambda s} r\,\mathrm{d}r\,\mathrm{d}\xi \qquad (\mathrm{RR}' \ll S).$$

The phase of this wavelet will depend upon the phase of the emitted wavelet, and on the optical path length between emitting point and the point under consideration. It is not important to know the absolute phase angle relative to that of the original wave surface, but it is often necessary to know the relative phase angle across the receiving surface. Usually, a convenient standard is taken; in the present example, this might be the phase angle on the axis in the O′R′ plane (i.e. at O′). In this case, in working out the relative phases of the arriving wavelets, the distance OO′ is subtracted from all paths. The result is termed the path difference Δ, and the resulting relative phase difference is, of course, $2\pi\Delta/\lambda$. Thus, for the wavelet arriving at the point R′ from the element $r\,\mathrm{d}r\,\mathrm{d}\xi$, the path difference is given by $\Delta=(\mathrm{RR}'-\mathrm{OO}')$, which can easily be derived from geometrical considerations, bearing in mind that the radius vector $\vec{\mathrm{OR}}$ is not necessarily orientated at the same angle as $\mathrm{O}'\vec{\mathrm{R}}'$.

The relative phase angle for the received wavelet is then $(\eta_0+2\pi\Delta/\lambda)$, where η_0 is the phase angle at (r, ξ) relative to that at the point O.

To integrate all the wavelets arriving at R′ from all points on the wave in the OR plane, while taking into account the relative phase, the wavelets are resolved into two components in quadrature phase relationship, and expressed as a complex quantity:

$$\delta\psi = \frac{\psi_0}{\lambda s}\left[\cos\left(\eta_0+\frac{2\pi\Delta}{\lambda}\right)+\mathrm{i}\sin\left(\eta_0+\frac{2\pi\Delta}{\lambda}\right)\right] r\,\mathrm{d}r\,\mathrm{d}\xi$$

or, using the exponential formulation:

$$\delta\psi = \frac{\psi_0}{\lambda s}\exp\left[\mathrm{i}\left(\eta_0+\frac{2\pi\Delta}{\lambda}\right)\right] r\,\mathrm{d}r\,\mathrm{d}\xi.$$

The total wave amplitude at R′ is then given by:

$$\psi' = \int_{-\pi}^{+\pi}\int_{0}^{\infty}\frac{\psi_0}{\lambda s}\exp\left[\mathrm{i}\left(\eta_0+\frac{2\pi\Delta}{\lambda}\right)\right] r\,\mathrm{d}r\,\mathrm{d}\xi. \qquad (2.1)$$

It will be noted here that ψ_0 and η_0 may be functions of r and ξ. If ψ_0 and η_0 are constant, the transmitted wave is

plane and of uniform amplitude. If η_0 only is constant, the wave is plane. In either case, the plane OR is a wavefront. If η_0 is not constant, the plane is, of course, not a wavefront.

Propagation of Wave Through a Microscope System

In calculating the propagation of waves through a microscope system, certain simplifications result if the calculation is carried out in stages and certain simplifying 'tricks' are applied. It will be appreciated already that the main problem in calculating wave propagation, apart from solving the final double integral, is the determination of the path difference to give the correct phase relationships. The process is simplified if a modification is introduced to divide the propagated wave into two components. To introduce this modification, it will be assumed for the present that the object is non-absorbing and that the path difference due to the changing refractive index is a function of the radius in the object only, that is to say, the object is symmetrical about the axis. It will also be assumed that no aberrations are present. These simplifying assumptions in no way limit the following argument, and are introduced solely to simplify the explanation. Departure from these conditions can readily be introduced if necessary, but only the axially symmetrical case will be used later in this monograph.

The distribution of amplitude and phase over a surface can clearly be represented by a distribution of vectors. In Figure 2.3 such a radial vector distribution is shown

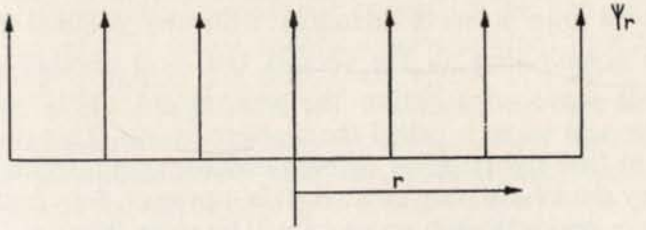


FIG. 2.3. *Vectorial representation of a plane wave.*

representing the plane-illuminating wave. On passing through the axially symmetrical phase object, these vectors will be rotated through an angle corresponding to the phase difference angle (η_r) given by:

$$\eta_r = 2\pi\Delta_r/\lambda$$

where Δ_r is the path difference introduced by the object at a radius r. The new distribution is represented by the vectors of Figure 2.4. The amplitudes are, of course,

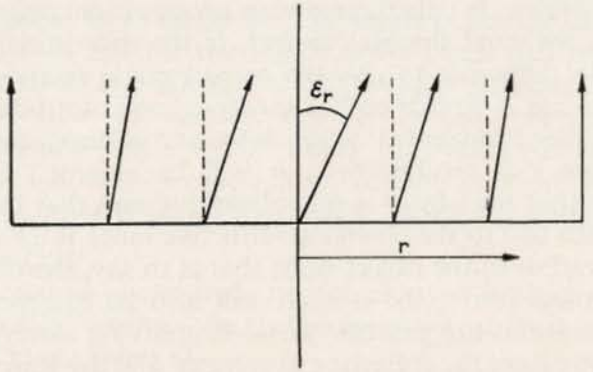


FIG. 2.4. *Vectorial representation of a plane wave perturbed by a phase object.*

unchanged since there is no absorption. In Figure 2.5 $\vec{OA}$ represents one of these perturbed vectors and $\vec{OB}$ the original unperturbed vector. The perturbed vector can be represented by the addition of the unperturbed vector and a joining vector $\vec{BA}$. If this procedure is adopted at every point, it will be seen that the perturbed wave has been resolved into a wave identical with the original plane wave constituting all the vectors $\vec{OB}$, and another component wave constituting the vectors $\vec{BA}$. This second component wave is called the scattered wave. This process may at first sight appear rather dubious, as it appears that energy must have been created. It is, however, fully justified in wave optics though no proof will be given here.

The propagation of the two component waves can now

be considered separately, and the resulting waves in the image plane compounded. It is convenient to consider the propagation in two stages, from the object to the back

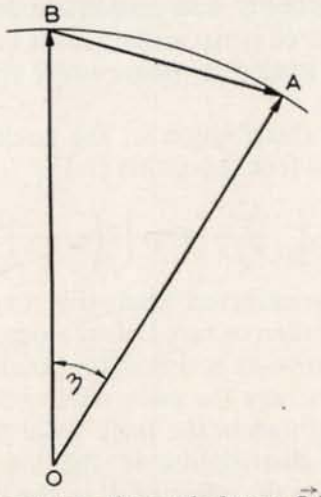


FIG. 2.5. *Vectorial resolution of perturbed wave* $\vec{OA}$ *into the unperturbed wave* $\vec{OB}$, *and the scattered wave* $\vec{BA}$.

focal plane of the objective lens and thence to the image plane. If, for the second stage, the waves are considered as reversing in direction of propagation and allowed to travel back into the object plane, the resulting distribution corresponds to that in the image, but in the usual optical manner, referred back to the object plane. The distribution can then be enlarged M times to give the true distribution in the magnified image plane, the *amplitude* being reduced at all points M times.

The plane wave component travels to the lens, and is there transferred into a spherical wave centred at a point on the axis and in the back focal plane. The wave then returns to the image plane with virtually no interference from any aperture in the back focal plane, since most of its energy is confined very near the axis. It can readily be shown that the effect is to give a plane wave in the image plane of equal amplitude to the original. In other words, this wave is unaffected by the aperture or the imaging process.

The scattered wave will have an amplitude equal to $2\psi \sin \eta_r/2$, or, if η_r is everywhere small, as is often the case in electron microscope specimens, and if the illuminating wave intensity and amplitude are taken as equal to *unity*, the scattered wave amplitude is equal to η_r, which, as we have seen, is also its phase angle with respect to the background wave.

The amplitude distribution in the back focal plane (ψ') can now be written from equation (2.1).

$$\psi' = \int_{-\pi}^{+\pi} \int_0^{\infty} \frac{\psi_r \eta_r \, r}{f\lambda} \exp\left[2\left(\eta_r + \frac{2\pi\Delta}{\lambda}\right)\right] \mathrm{d}r \, \mathrm{d}\xi \qquad (2.2)$$

It must be remembered that this expression is only accurate if η_r has fallen to zero before r becomes comparable to f, a condition always satisfied for small electron microscope objects. Δ is here the path difference between points in the object and those in the back focal plane.

The amplitude distribution in the image plane (ψ'') is given by integrating the effect of ψ' in the image plane:

$$\psi'' = \int_{-\pi}^{+\pi} \int_0^{r_0'} \frac{\psi' r'}{f\lambda} \exp\left(\frac{2\pi \mathrm{i}}{\lambda} \Delta'\right) \mathrm{d}r' \, \mathrm{d}\xi \qquad (2.3)$$

where ψ' is the solution of equation (2.2) and r_0' is the radius of the physical aperture in the objective lens. Equations (2.2) and (2.3) are the well-known Fourier integrals, and their solution allows the intensity distribution in the image of an object to be calculated. This will be given by the square of the sum of the background wave amplitude, which by definition is unity, and that of the scattered wave (ψ''):

$$I = (1 + \psi'')^2$$

or, if $\psi'' \ll 1$

$$I = 1 + 2\psi''. \qquad (2.4)$$

Consider the system shown in the diagram of Figure 2.6. OR is the object/image plane, OB the axis, and BR′ the back focal plane. LL′ is the lens. The distance OL and LB are both equal to f, the focal length of the lens.

A wavelet initiated at a point P, (r, ξ) in the object plane expands spherically outwards towards the lens, as indicated by the dotted lines. On traversing the lens, this wavelet is transformed into a plane wave tilted with respect to a

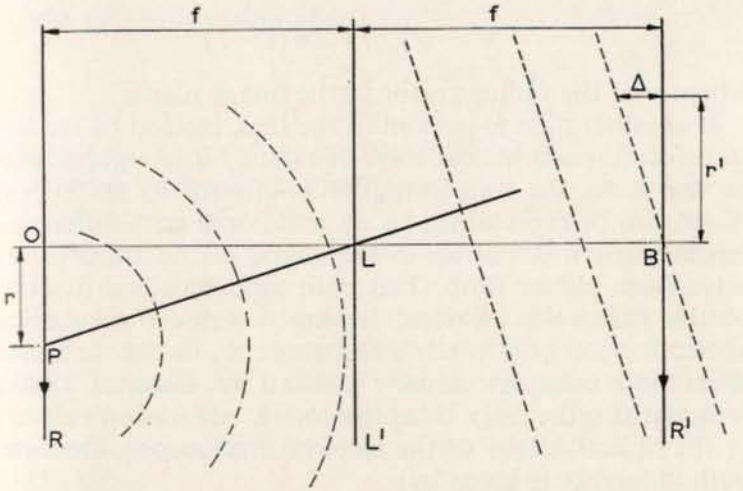


FIG. 2.6. *Diagram illustrating the derivation of path difference for a wavelet propagated through the objective lens.*

plane normal to the lens axis by an angle given by r/f (where $r \ll f$), this being the effect of an aberration-free lens. The plane wavelet will have an amplitude $\psi_0/f\lambda$, and this will not change between the lens and the back focal plane, since a uniform plane wave is propagated with constant amplitude. On arriving at the point B, the axial point in the back focal plane, a reference surface for the path difference determination is established by the wavefront. All points on this bear a constant phase relation to a wavelet leaving the axial object point. Hence, the path difference is simply the distance from the wavefront of the wavelet to the point in the back focal plane being considered. It is easily seen that for a point of co-ordinate position $(r'\xi')$, this difference is given by:

$$\Delta = -(r'r/f) \cos(\xi - \xi').$$

The negative sign occurs where the reference wavefront has already crossed the point under consideration, as it has here, when r and r' are of the same sign.

By similar argument the path difference from back focal plane to image is:

$$\Delta = -(r'r''/f)\cos(\xi - \xi')$$

where r'' is the radius vector in the image plane.

If an aberration is present in the lens, instead of the lens transforming a spherical wave of radius f into a plane wave or vice versa, the transformation is in error by an amount which can be represented by an additional path difference, representing a deviation in the shape of the transformed wave from planar form. This path difference is a function of the radius in the lens, having a value for spherical aberration of $\frac{1}{4}C_s r^4/f^4$ ($r \ll f$), where C_s is the spherical aberration constant already defined in Chapter I. For astigmatism, the only other geometric aberration relevant in the objective lens of the electron microscope, the extra path difference is given by:

$$\Delta = \tfrac{1}{4}(z_a r'^2/f^2)\sin\xi/2,$$

where ξ is the orientation angle relative to the main astigmatic plane.

Usually the analytical solution of the imaging integrals is not possible, and it is necessary to resort to numerical methods. The elementary introduction to wave-optical image formation is given here, as it is necessary for the derivation of the contrast conditions in the electron microscope (Chapter IV); and an understanding of it is highly desirable for anyone wishing to appreciate fully the functioning of the instrument. As an example, the theory will be applied to one simple but very representative case, a small phase object (no absorption), of circular shape and radius r_0, and introducing a uniform small phase lag η_0 into the illuminating beam.

It will readily be seen that the amplitude of the scattered wave is given by η_0 and that it will have a phase angle

$\cos^{-1}\eta_0/2$ relative to the background. The amplitude of the wave in the back focal plane is given by:

$$\psi' = \int_{-\pi}^{+\pi}\int_0^{r_0}\frac{\eta_0 r}{f\lambda}\exp\left(\frac{2\pi i r r'}{f\lambda}\right)\cos\xi\,dr\,d\xi.$$

The first integration gives:

$$\psi' = \int_0^{r_0}\frac{2\pi\eta_r r}{f\lambda}J_0\left(\frac{2\pi r r'}{f\lambda}\right)dr$$

where J_0 is the zeroth order Bessel function.

A second integration gives:

$$\psi' = \frac{\eta_r r_0}{r'}J_1\left(\frac{2\pi r_0 r'}{f\lambda}\right)$$

This distribution is that of the well-known Airy pattern $[J_1(x)/x]$. It will be noted that the microscope resolving power is given by $0{\cdot}61\lambda/\alpha_0 = 0{\cdot}61\lambda f/r_0'$, where r_0' is the radius of the objective aperture. If $r_0 \ll d$, the Bessel term argument is small and $J_1(2\pi r_0 r_0'/f\lambda) \approx 2\pi r_0 r_0'/f\lambda$.

Hence:

$$\psi' = \frac{\pi\eta_r {r_0}^2}{f\lambda} \tag{2.5}$$

or the amplitude is constant across the aperture. This is the case which will be considered hereafter.

The amplitude in the image plane is given by:

$$\psi'' = \int_{-\pi}^{+\pi}\int_0^{r_0'}\frac{\psi' r'}{f\lambda}\exp\frac{2\pi i r' r''}{f\lambda}\sin\xi\,dr'\,d\xi$$

where r'' is the radius vector in the image plane, and r_0' is the radius of the lens aperture.

The first integration gives:

$$\psi'' = \int_0^{r_0'}\frac{2\pi\psi' r'}{f\lambda}J_0\left(\frac{2\pi r' r''}{f\lambda}\right)dr' \tag{2.6a}$$

and the second

$$\psi'' = \frac{\psi' r'}{r''}J_1\left(\frac{2\pi r_0' r''}{f\lambda}\right).$$

Putting in the value of ψ' from equation (2.5)

$$\psi'' = \frac{2\pi^2 r_0^2 r_0'^2 \eta_r}{(f\lambda)^2} \frac{J_1(2\pi r_0' r''/f\lambda)}{(2\pi r_0' r''/f\lambda)}. \qquad (2.6b)$$

Putting $r_0/f = \alpha_0$, the objective aperture angle, and using the expression $\alpha = 0{\cdot}61\lambda/\alpha_0$ for the microscope resolution, we obtain:

$$\psi'' = 7{\cdot}3 \frac{(r_0/\alpha)^2 \eta_r J_1(1{\cdot}22\pi r''/\alpha)}{1{\cdot}22\pi r''/\alpha}.$$

It will be of value to examine the above result a little further. In order to obtain the intensity distribution in the image plane, the scattered wave amplitude must be added to the background. To do this, the relative phase between the scattered wave and background wave is required. The fact that the imaginary component has disappeared in the integrations, indicates that the scattered wave has the same relative phase in the image as at the object. Figure 2.7 shows the vector relationships: $\vec{OB}$ is

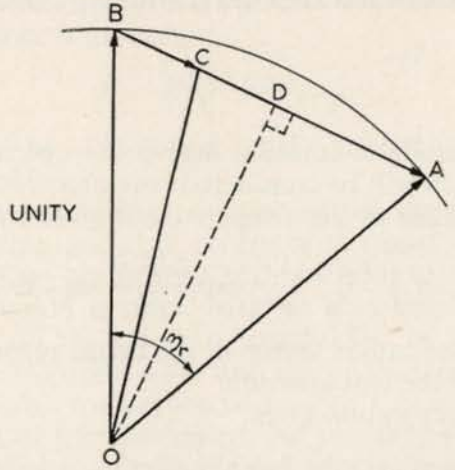


FIG. 2.7. *The scattered wave in the image plane* $\vec{OC}$ *is compounded with the background wave* $\vec{OB}$ *to give the total wave amplitude* $\vec{BC}$.

the background wave, $\vec{OA}$ the phase delayed wave, $\vec{BA}$ the scattered wave and $\vec{BC}$ the scattered wave in the image.

OD is the perpendicular bisector of BA. The total wave intensity (amplitude squared) at the image is given by $(\vec{OC})^2$ where:

$$\begin{aligned} \mathscr{I} = (\vec{OC})^2 &= (\vec{OD})^2 + (\vec{CD})^2 \\ &= (\vec{OB})^2 - (\vec{BD})^2 + (\vec{BD} - \vec{BC})^2 \\ &= (\vec{OB})^2 + (\vec{BC})^2 - 2(\vec{BD})(\vec{BC}) \end{aligned}$$

If $\eta_r \ll 1$, $\vec{BD} = \eta_r/2$, hence:

$$\mathscr{I} = 1 + (\psi'')^2 - 2\psi''\eta_r/2.$$

Since $r_0 \ll d$, $\psi'' \ll \eta_r$, and hence:

$$\mathscr{I} = 1 - \psi''\eta_r. \qquad (2.7)$$

This is the image intensity, and the contrast against the unit intensity background is therefore given by $\psi''\eta_r$. It will be noted that the distribution of contrast is as for the Airy *amplitude* pattern and not the Airy *intensity* pattern as obtained for a small radiating source. The distribution is shown in Figure 2.8.

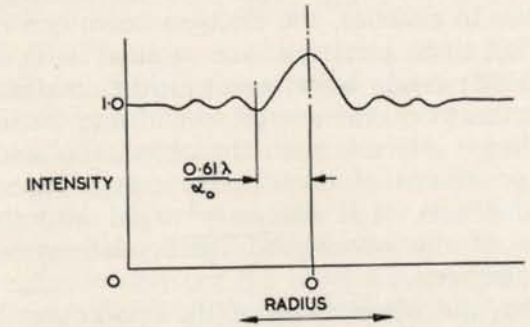


FIG. 2.8. *The intensity distribution produced by a small phase object.*

If the scattered wave were changed in phase by about $\pi/2$ it would now be in phase or anti-phase with the background, and the intensity would become:

$$\mathscr{I}_p = 1 \pm \psi''.$$

Since η_r is small by definition, $\mathscr{I}_p$ is larger than $\mathscr{I}$, and can

be very much larger. This is the basis for phase contrast, and will be discussed further in Chapter IV. In fact, when η_r is very small, as it may often be in the electron microscope, the distribution in the Gaussian focal plane means very little, since a very small defocusing, which has the effect of rotating the scattered wave vector with respect to the background vector, can make a very large contrast change, even without defocusing enough to spoil the resolution.

With this simple example, the theory of wave propagation will be left until it is required in later chapters.

It is again stressed that the simple example is chosen to indicate the method more clearly. The extension to more complicated objects, and to include aberrations and departures from perfect focus, complicates the mathematics but does not essentially change the method.

Emission, Waves, Coherence

The wave theory has, so far, been discussed as if the waves were continuous in time and the electron beam monokinetic. In practice, the electron beam can never be monokinetic, since electrons are emitted with a finite energy spread; and, after acceleration, maintain this spread. Furthermore, the energy spread may be increased when the beam interacts with the object. To understand how eiectron beams of non-discrete energy can produce interference effects, it is necessary to go into the basic significance of the waves and their relations with the ndividual electrons.

Essentially, the whole basis of the connection between particle and wave is the uncertainty relation. We have a stream of particles emitted with random energy and direction. Thus, for any particular particle, for example the next one to be emitted after the present moment, we cannot prescribe the energy or direction with certainty; but we can say that, since the *distribution* of energies and direction are known, we can prescribe the probability of the energy and direction falling within any specific limits.

Thus the most probable emission energy is kT, where k is Boltzmann's constant and T the absolute temperature, and the distribution is Maxwellian.

Now the equivalent wave will have properties which represent these uncertainties, so that calculations based on the waves will only give the probability of any particular electron falling within any prescribed area, but will give the actual distribution for a large number of emitted electrons, the intensity at any point being the square of the amplitude at that point.

If an infinite sine wave of amplitude ψ_1, frequency ν, velocity of propagation v_ϕ, and wavelength λ is associated with an electron of specific energy $\mathscr{E}$, the position of the electron at any moment of time is completely undefined. This brings us to the very basis of the 'uncertainty principle' which says that the closer we can define the momentum of a particle, the less closely can we define its position. The uncertainty in momentum component (δM_x) is related to the uncertainty in position (δx) by the well-known expression of Heisenberg:

$$\delta x . \delta M_x = h$$

where h is Planck's constant ($=6{\cdot}624 \times 10^{-27}$ erg sec). Rather than start from this relation and derive the particle wave relationship, we shall proceed from a different direction and show that it leads to this result. Suppose we represent each elemental spread in the energy uncertainty of the beam by an infinite (or near infinite) wave of frequency ν, wavelength λ and velocity $v_\phi = \lambda\nu$.

de Broglie, in deriving the wave mechanical hypothesis, started from the well-known relationship between the frequency of a light wave and its energy:

$$\mathscr{E} = h\nu. \tag{2.8}$$

The energy of a particle with rest mass m (unlike the photon with zero rest mass) is given by:

$$\mathscr{E} = mc^2/(1 - v^2/c^2)^{1/2}$$

where v is the particle velocity.

It will be seen that if the relativistic term in the denominator is expanded by the binomial theorem:

$$\mathscr{E} \approx mc^2 + \tfrac{1}{2}mv^2$$

which is the sum of the mass energy and the kinetic energy.

Putting this value of energy in equation (2.8) gives:

$$\nu = mc^2/h(1 - v^2/c^2)^{1/2}. \tag{2.9}$$

We now give the relativistic value of wavelength as:

$$\lambda = (h/mv)(1 - v^2/c^2)^{1/2}. \tag{2.10}$$

This is given without proof or justification; it will be remembered that it is part of a hypothesis which has proved consistent with experimental findings.

The propagation velocity v_ϕ $(=\lambda\nu)$ of the electron wave is now given from (2.9) and (2.10) as:

$$v_\phi = c^2/v. \tag{2.11}$$

It will be noted that this velocity is inversely dependent on the mean particle velocity, and is greater than that of light. This does not concern us since it is a phase velocity and does not imply the propagation of energy at a velocity greater than that of light.

Having established these hypothetical expressions, it is now possible to derive the wave system which represents a given electron beam. Suppose a beam is produced by accelerating electrons to an electrostatic potential ϕ_0 from an infinite cathode, with an emission velocity distribution in the direction of acceleration $f(v_e)$, so that the number of electrons emitted from unit area between a velocity v_e and $v_e + \delta v_e$ is given by:

$$\delta N = N_0 f(v_e)\, \delta v_e$$

where N_0 is the total number of emitted electrons per unit area or the intensity. The amplitude of the contribution due to these electrons is $(\delta N)^{1/2}$.

If δv_e is small, these electrons will be represented by a near-infinite wave having values of wavelength, frequency, and velocity given by equations (2.10), (2.9), (2.11),

where v is given the value $\sqrt{(2e\phi_0/m)} + v_e$, $\sqrt{(2e\phi_0/m)}$ being the velocity acquired during acceleration.

It will be noted that we are now considering the wave to represent a large number of electrons. This is convenient at present, although, as has been said before, it is later more convincing to associate the wave packets with individual electrons.

To obtain the effect of all electrons, we must clearly integrate the near-infinite waves due to the elemental groups. This procedure means the adding of a series of near-infinite waves whose frequency and wavelength vary over a range corresponding to the variation of particle velocity, the waves all being in phase at the moment of emission.

Clearly, such a series of waves of slowly varying wavelength can only be in phase at one position in the course of propagation. Away from this point, the phases will become more and more dispersed, and will eventually be so randomised that the sum of the amplitudes will be zero. Thus, the effect of the spread in momentum or wavelength is to restrict the total wave to a discrete length in space and time. It may easily be seen that the mean length of the wave will correspond to that at which the wavelength difference corresponding to the mean velocity spread is one wavelength.

Now:

$$\lambda = h/mv.$$

Differentiating:

$$\delta\lambda = -\lambda \frac{\delta v}{v}.$$

If there are n wavelengths in the mean length of the wave train (δx), and δv is put equal to the mean velocity spread of the emitted electrons, then the length is obtained by putting $n\delta\lambda = \lambda$; so,

$$n\delta\lambda = n\lambda \frac{\delta v}{v} = \lambda$$

or, since $n\lambda = \delta x$

$$\delta x = \frac{\lambda v}{\delta v}$$

but

$$\lambda = h/mv$$

Hence

$$\delta x = \frac{h}{mv} \times \frac{v}{\delta v}$$

$$= h/\delta mv$$

or

$$\delta x \delta mv = h \tag{2.12}$$

which is the 'uncertainty principle'.

The velocity of the wave packet as a whole has not so far been specifically derived. We have seen how the packet is made up by the superposition of infinite waves of discrete energy, of varying phase velocity, and wavelength. Again, we can best appreciate the propagation of the wave packet in a qualitative manner. If all the infinite waves are in phase at one point at one instant, then a little later the peaks which were in phase will have moved by different amounts. Suppose we consider the two infinite waves corresponding with zero emission velocity, and with the mean emission velocity. The wavelength of either of these is given by:

$$\lambda = \frac{h}{mv}\left(1 - \frac{v^2}{c^2}\right)^{1/2}.$$

Differentiating:

$$\frac{\delta\lambda}{\lambda} = -\left(\frac{c^2}{c^2 - v^2}\right)\frac{\delta v}{v}. \tag{2.13}$$

$\delta\lambda/\lambda$ gives the fractional difference in wavelength between the two infinite waves if δv is put equal to the mean emission velocity. Similarly, the difference in phase velocity is given by the differentiation of:

$$v_\phi = c^2/v$$

$$\delta v_\phi = -\frac{c^2}{v}\,\delta v/v. \tag{2.14}$$

Now, it will readily be seen that, due to the 'vernier' action of the waves sliding one over the other, if one wave moves a certain distance relative to the other, the point of phase coincidence will move a distance $(\lambda/\delta\lambda)$ times this distance, where again $\delta\lambda$ is the small difference in wavelength. Thus, the velocity of the point of phase coincidence relative to the waves is $(\lambda/\delta\lambda)$ times the relative phase velocity (δv_ϕ). The absolute velocity of the point of phase coincidence (v_g) is therefore given by:

$$v_g = v_\phi - (\lambda/\delta\lambda)\,\delta v_\phi.$$

Putting in values of $(\lambda/\delta\lambda)$ from equation (2.13) and δv_ϕ from equation (2.14), we get:

$$v_g = \frac{c^2}{v} - \frac{c^2}{v}\left(\frac{c^2 - v^2}{c^2}\right) = v.$$

The wave packet is now seen to proceed with the electron, i.e. with the velocity of the electron. Its significance may be regarded thus: imagine any one of the emitted electrons chosen at random. One could not say in advance what emission energy it could have, except insofar as previous experience could predict the probability of it having any particular energy since the distribution of emission velocities is known. The wave packet proceeds with the mean velocity and its envelope represents the distribution of emission energies. The wave will, of course, propagate outward over a volume of space much larger than that encompassed by the trajectory of any one electron, but equal to that containing the trajectories of all electrons. The amplitude of the wave at any point gives the square root of the probability of any particular electron crossing that point. Thus when the electron waves interact to form an interference pattern, the distribution of intensity, or the square of the amplitude, gives the probability of any one electron falling at any point in this distribution. For a large number of emitted electrons, therefore, the distribution of intensity of the wave gives the distribution of the electrons.

Coherence

The term 'coherence' has been the source of much misunderstanding by students. Difficulty comes about because the term is used to explain an effect arising from several different causes.

Consider first in a qualitative manner, the question of interference between and within wave packets. By interference, we mean interaction in a manner where relative phases are of importance. If n identical packets all fall on a screen at exactly the same time and in identical phase, the resulting amplitude will be n times that of one packet. The intensity will therefore be n^2 times that of one packet. If these packets fall at separate times so that no overlap in arrival time occurs, there can be no interaction and the *intensity* will be n times that due to each packet. Now what happens when the arrival times overlap, as they may well do in practice? Any two wave packets may overlap in time and will then add in effect according to their relative phase —that is, giving an intensity between zero and four times that due to one, depending on their relative phase relationships and the degree of overlap in time. However, since arrivals (and phase differences) are random in their incidence, the net effect is just as if no interaction occurred, since the sum of n sine waves of random phase, though identical in frequency, is proportional to $\sqrt{n}$. Hence it is always considered that no interference can occur between one wave packet and another. This is not strictly true, but it is effectively so, since the result is no different from what would occur if no interaction took place.

We now consider the effect of two parts of a given wave packet interacting. This will happen, for example, when a wave packet is split up in some way and then the separate parts are recombined. A simple example is the Young's interference experiment. A wave falls on a barrier pierced by two small apertures. The parts of the wave transmitted through the apertures then spread out and overlap. When they fall on a screen, interaction occurs; and, since the parts originate from a particular wave packet, they will

have a definite phase relation, depending on the total path difference, taken from the source through the apertures and up to the point under consideration. Furthermore, all wave packets *arising from one point* on the source will have this same phase relation between their individual parts, and hence will produce the same amplitude pattern, i.e. an interference or diffraction pattern.

If the path difference between the separate parts of the wave packet is larger than the total length of the packets, clearly the two parts cannot interact since the two parts in effect arrive at different times. For shorter path differences, the two parts may only partly overlap in time, and only the overlapping part of the packets will interact. This effect is seldom of significance in practice, for the electron wave packet is of the order of 10^5 wavelengths long.

Suppose now that one part of a wave is scattered by interaction with the fields of an atomic array, such that an increase in the spread of energy occurs. The effect of this will be to reduce the length of the wave packet, since the spread in momentum is increased. In addition, the wavelength and the frequency will be modified. Consider an experiment where a wave packet initiating from a source is split into two parts, and one part passes through an object and is later brought together with the second part. The object is assumed to extract some energy from the beam, i.e. to increase the energy spread. The two parts of the wave packet are split into two packets of different wavelength and frequency, but with a definite relation between their times of arrival. The scattered packet will be shorter than the unscattered one. It must be noted that we are not concerned with the relative distribution of amplitude in space, but rather this distribution in time, since it is the relative phases at the moment of arrival which is relevant to the effect at the fluorescent screen, for example.

Consider the two parts of the wave packet, one having a mean velocity $v_0 + v_1$ and the other $v_0 + v_2$, when v_1 and v_2 are the mean velocity spreads for the two parts.

Consider the two parts of the wave packet as represented by (a) and (b) in Figure 2.9: (b) is shorter than (a) because

it has a larger uncertainty in momentum. The frequencies and velocities are also different. Now, consider the figure to represent the variation of amplitude with time rather than velocity. It is clear that interference can only take

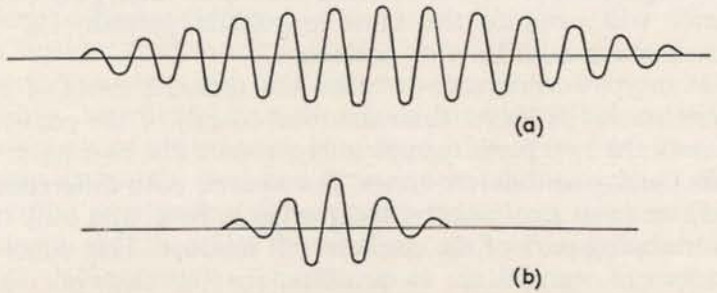


FIG. 2.9. (a) *Unscattered and* (b) *scattered wave packets.*

place over a time corresponding to the duration of the shorter packet and only then if the relative phase change between the packets is small (e.g. < half a period).

The periodic time of the oscillation can be written:

$$\tau = \frac{1}{\nu} = \frac{h(1 - v^2/c^2)^{1/2}}{mc^2}.$$

Differentiating:

$$\delta\tau = \frac{h}{mc^2} \frac{v/c^2}{(1 - v^2/c^2)^{1/2}} \delta v.$$

If the wave packet contains n periods, and $\delta\tau$ is the difference in period between the two wave packets, $n \times \delta\tau$ must be small compared with τ for effective interference to occur.

For the shortened wave packet:

$$n = \delta x/\lambda$$
$$= h/\lambda\delta mv \quad \text{(from equation (2.12).}$$

Hence:

$$\frac{h}{\lambda\delta mv} \times \frac{h\delta v}{mc^2} \times \frac{v/c^2}{(1 - v^2/c^2)^{1/2}} < \frac{h(1 - v^2/c^2)^{1/2}}{mc^2}$$

but

$$\delta mv = m\delta v.$$

Hence:

$$v^2/c^2 < 1 - v^2/c^2$$
$$v < 0{\cdot}7c.$$

Thus only for velocities well in the relativistic region is there any restriction to stop interference between an inelastically scattered part of a wave packet and the unscattered part.

We have so far considered coherence in terms of wave packets starting from one point on the emitting source, and have shown what effects can limit the formation of interference fringes or patterns. There is one more effect which should be considered in this category. In electron optics, the electrons are accelerated by the application of an accelerating potential. Since this potential can never be perfectly constant, the electron wavelength and the wave packet velocity will vary as the potential varies. The latter variation results in a variation in optical path, measured in wavelengths, and hence relative phase values vary. The result of this is that the size of a resulting diffraction, or interference, pattern varies and, at some points, fringes will overlap and tend to cancel out. The effect may be termed 'chromatic incoherence', and is discussed for example in relation to Fresnel fringes in Chapter V.

In addition to these effects which are pertinent to a point emission source, the extent of an interference pattern is limited by the finite size of the emission source which occurs in practice. This effect can be evaluated quite simply after calculating the diffraction pattern which would be produced by a point source, by considering the effect of moving the source laterally by a distance equal to the size of the actual source. In general, but not necessarily always, the diffraction pattern will be displaced by a proportionate amount. Diffraction fringes narrower than the movement of the pattern will be wiped out, and those wider will suffer a loss of intensity.

Coherence in the Microscope

Although it is possible to make generalisations concerning the question of coherence in the microscope, this tends

to hinder rather than help the understanding of the physical significance of coherence.

In the simplest instance, the microscope object is self-radiating, and the radiated waves from different points are emitted at random and are, therefore, incapable of regular interference. The emitted radiation can then be said to be incoherent. It must be remembered, however, that each wave is within itself coherent, and passes the lens to converge in the image plane as a coherent whole to produce an Airy distribution. The separate distributions from separate points are incoherent, and the intensity produced at any point is the linear sum of the squares of the amplitudes in all the Airy patterns corresponding to all the emitting points:

$$\mathscr{I}_t = \psi_i^2 = \sum \psi^2_{xy}.$$

If, on the other hand, the object is illuminated from a very small source, then waves originating from the source fall all over the object, producing coherent scattered waves from the scattering points in it. These waves diverge into the lens and thence converge to a focus in the image to produce Airy patterns of amplitude ψ_{xy}, where x, y may be the co-ordinates from the Gaussian image points corresponding to the scattering points in the object. The image intensity is now the square of the vectorial sum of all the amplitudes due to the different Airy distributions resulting from the scattering points:

$$\mathscr{I} = \psi_i^2 = (\sum \psi_{xy})^2.$$

One example of the difference between the image produced under incoherent and coherent illumination is that which occurs for two closely adjacent points (radiating or scattering). As these points are brought closer and closer together, the corresponding central maxima in the image will eventually merge to give a single maximum. For the incoherent case, the maxima still have a 18% dip between the intensity maxima, when the spacing is equal to the microscope resolving power, or the radius of the Airy central maximum. For the coherent case, the same is only

true for a spacing twice as big. The points would thus be more easily distinguishable in the incoherent case.

So far, two extremes have been described. If now a source of finite size is considered, an intermediate situation can result. Consider a system in which a source of radius r_0 (Figure 2.10) illuminates an opaque disc containing two very

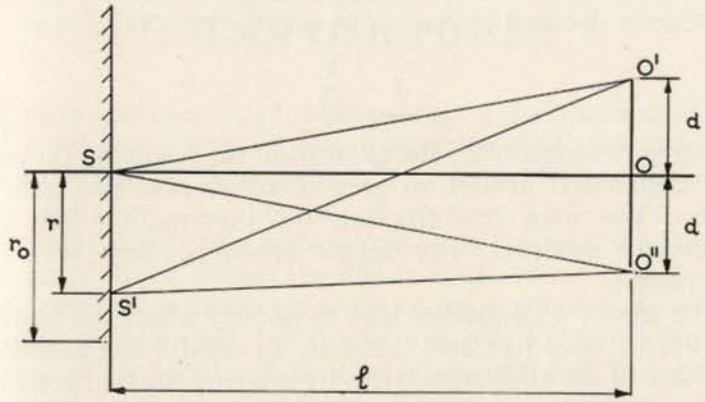


FIG. 2.10. *Diagram illustrating the incoherence produced by finite source size.*

small apertures, a_1 and a_2, separated by a distance $2d$, placed in the object plane of a microscope. Each point in the source will irradiate the apertures; these, if very small, will produce spherical waves which will pass through the microscope to converge into the image plane, there to give two Airy amplitude distributions which will add vectorially. Any other point on the source distinct from the first will act likewise, but the spherical waves scattered by the apertures will have a phase difference differing from those for the first source point, by an angle given by $2\pi rd/l\lambda$, where l is the source-object distance. It follows that this phase difference occurs between the Airy amplitudes at any point in the image plane. Thus, considering the whole source, phase differences over all values from 0 to $2\pi r_0 d/l\lambda$ will exist between the Airy amplitude patterns. If this range exceeds $\pi/2$, then the coherence is effectively destroyed; but, unlike the Young's diffraction pattern

previously considered, the pattern is not wiped out since the maxima are not displaced in position. The net result, if r_0 becomes very large, is to give exactly the same pattern as would be obtained if the apertures were self radiating, or incoherent. The condition is approached when $2\pi r_0 d/l\lambda \to \pi$ or $r_0 \to l\lambda/2d$.

It will be noted that, if the semi-angle subtended by the cathode is denoted by $\theta = r_0/l$ (θ small), the condition is:

$$d \sim \frac{\lambda}{2\theta}$$

which is reminiscent of the expression for resolving power. The distance d defines an area of coherence. For points spaced less than this amount, the illumination will be effectively coherent. For larger spacings, they will be incoherent.

The use of a condenser lens in no way alters the above argument, since its effect is simply to control the position and size of the effective source, by imaging the real source.

CHAPTER III

THEORETICAL LIMITATIONS TO RESOLVING POWER

THERE has been much controversy on the question of the resolving power of the microscope, both optical and electron optical. Most of this has resulted from differences of definition, the source of most controversies! The resolving power of a microscope is, as the term suggests, a quantity which defines the ability of the microscope to see fine detail. It is defined in terms of the smallest detail which can be made visible, but this very clearly will vary with the shape of the detail and its other properties, such as contrast relative to its surroundings. The practical microscopist requires a definition which he can relate to something he can observe on his instrument, or he may require a simple test to determine the resolving power of a given instrument. The theoretician concerned with the microscope physics will, on the other hand, give a definition related to some mathematical result which may not be directly connected with any easily realisable experiment. The question is greatly simplified if resolving power is considered in two parts, the first being a function of the instrument design, and the second a function of the object. Thus, in the first instance, the capabilities of the instrument may be considered in terms of an ideal and very simple object, and secondly, the limitation imposed by the object may be considered.

The simplest possible object is a single point, emitting monochromatic radiation. In an aberration-free system, such an object will result in a uniform wave amplitude distribution of constant phase in the back focal plane. The amplitude distribution in the image plane is then given

from equation (2.6) and the intensity at each point is, of course, the square of the amplitude. The solution is the well-known Airy intensity distribution which has the appearance shown in Figure 3.1 (curve 'a'), consisting of a

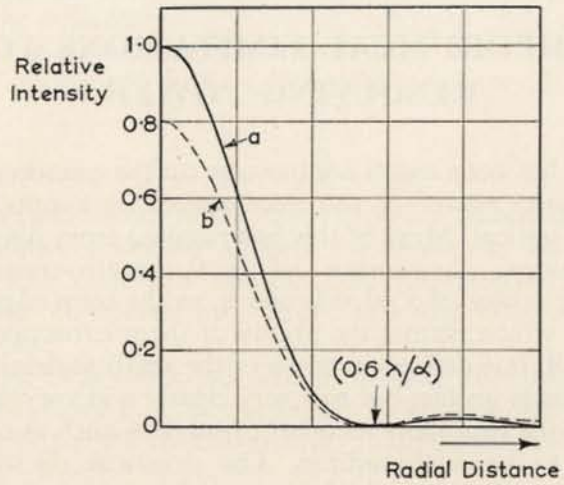


FIG. 3.1. *Comparison between the Airy diffraction pattern* (a) *and the equivalent distribution with one wavelength of spherical aberration* (b).

central maximum falling to zero at a radius given by $0{\cdot}61\lambda/\sin\alpha_0$, and thereafter a series of bright and dark rings of falling intensity. It is convenient to define the resolving power (d) as the radius of the central maximum of this diffraction pattern. This is strictly the resolving power as limited by diffraction.

If every point on the object emits spherical waves of varying amplitude which are incoherent with one another, and if the optical system is aberration-free, the image intensity distribution can be computed by a direct integration of an infinite array of Airy diffraction patterns of intensity. In such a case, the resolving power can be directly and simply related to some simple test such as a measurement of the sharpness of the image of a straight edge. If, on the other hand, the object is not self-radiating, and is illuminated with coherent radiation, as is usually the case

in electron microscopy, the application of a simple test, such as described, becomes much more arbitrary. This should always be borne in mind. If, by tests which will be described later, it has been shown that the resolving power of a given instrument, as previously defined, is 10 Å, this does not mean that all details down to 10 Å will be visible. The visibility of given detail depends not only on the resolving power but also on the coherence of the illumination, the aberration of the optical system, the shape and disposition of the detail, and the contrast conditions. In many cases detail larger than 10 Å may not be visible, and in some special cases, detail appreciably less than 10 Å may be visible. It is impossible in a small space to enlarge upon this very complex subject, but at least its importance must be stressed, while its implications are by-passed by defining the resolving power in terms of the radius of the central diffraction maximum in the image of a point object.

Unfortunately, this simple definition cannot be accepted without reserve. When the effects of aberrations are considered, it is found that though they modify the Airy distribution, they do not necessarily change significantly the shape of the central maximum, but rather transfer energy from this to the rings. By definition, they therefore would not change the resolving power so long as this is their only effect. In fact, by the time more than a certain percentage of the energy has been removed from the central maximum of the Airy pattern, it is no longer reasonable to consider the definition as holding. Possibly a further condition might be imposed wherein the resolving power is defined as the radius of the central diffraction maximum resulting from a point-emitting object when more than, say, 80% of the energy arriving in the image plane is contained therein. This still gives a somewhat arbitrary definition, but it at least gives a basis for comparison and discussion.

The Effect of Geometric Aberrations

The presence of aberrations in the objective lens can be calculated by introducing the equivalent phase changes into the imaging integrals. It has already been explained that a

perfect lens transforms spherical waves or wavelets radiating from the object into spherical waves or wavelets converging on to the image plane. Aberrations represent an error in this transformation and are represented by an additional path difference which is a function of the co-ordinates of the origin of the wavelet in the object plane and of those of the point considered in the aperture plane. In the electron microscope, only spherical aberration and astigmatism are of significance, and both of these are independent of the object plane co-ordinates. The path differences introduced are given respectively by:

Spherical aberration,

$$\Delta = \tfrac{1}{4}C_s\alpha^4 = \tfrac{1}{4}C_s(r'/f)^4. \qquad (3.1)$$

Astigmatism,

$$\Delta = \tfrac{1}{4}z_a\alpha^2(\sin\zeta/2) \qquad (\alpha \text{ small}). \qquad (3.2)$$

It is not necessary to consider astigmatism as imposing any fundamental limitation because it only arises as a result of geometrical imperfections in the lens and, as will be described in Chapter V, can be reduced to an arbitrarily small level. Thus it does not necessarily constitute a real limit to microscope performance.

In considering the effect of spherical aberration on the Airy distribution, it is necessary also to consider the focus as a variable parameter, since optimum results are not obtained in the true Gaussian image plane. It is well known that the plane of minimum diameter of the *geometric* caustic obtained in the presence of spherical aberration is two-thirds of the distance from the axial point where the paraxial rays cross the axis, to that point where the extreme marginal rays cross. The wave optical calculation shows this not to be the optimum plane. The focus parameter is introduced by the insertion of a further path difference given by:

$$\Delta = \tfrac{1}{2}\delta f(\alpha)^2 = \tfrac{1}{2}\delta f(r'/f)^2$$

where δf is the axial distance from the Gaussian focus.

The solution of the imaging integrals with the phase

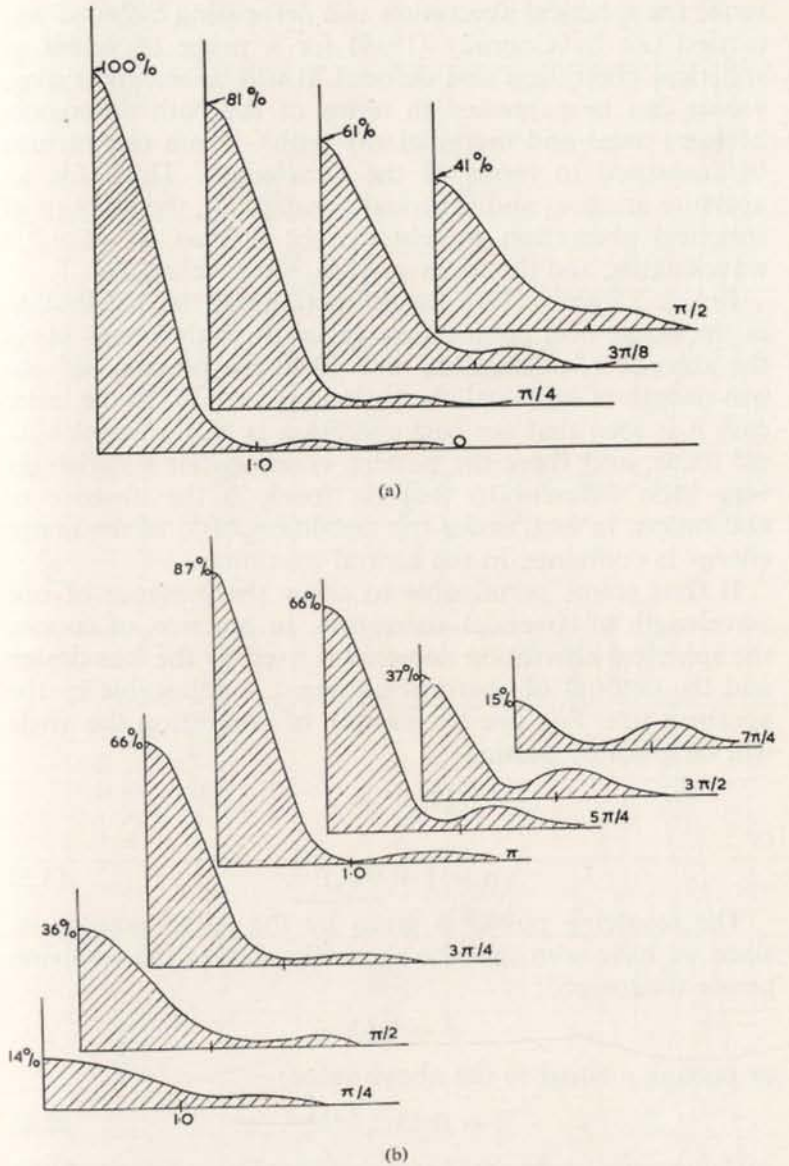


FIG. 3.2. *Variation of the intensity distribution from a point object with radius as a function of focus,* (a) *with no spherical aberration,* (b) *with one wavelength spherical aberration.*

terms for spherical aberration and defocusing included was carried out by Conrady (1919) for a range of values of spherical aberration and defocus. It will be seen that these values can be expressed in terms of the path differences between axial and marginal ray paths. These can in turn be measured in terms of the wavelength. Thus, for an aperture angle α_0 and spherical constant C_s, the amount of spherical aberration present can be defined as $\frac{1}{4}C_s\alpha_0^4/\lambda$ wavelengths, and the focus as $\frac{1}{2}\delta f\alpha_0^2/\lambda$ wavelengths.

Figure 3.2 shows how the diffraction pattern distribution in the image of a point object varies through focus—(a) in the absence of aberration, and (b) in the presence of one wavelength of spherical aberration ($\frac{1}{4}C_s\alpha^4 = \lambda$). In the latter case it is seen that the best condition is half a wavelength off focus, and there the pattern is somewhat surprisingly very little different to that on focus in the absence of aberration. In fact, under this condition, 80% of the image energy is contained in the central maximum.

It thus seems permissible to allow the presence of one wavelength of spherical aberration. In practice, of course, the spherical aberration constant is fixed by the lens design and the amount of aberration present is adjustable by the aperture size. For one wavelength of aberration the angle will be given by putting

$$\tfrac{1}{4}C_s\alpha^4 = \lambda$$

or

$$\alpha = 1{\cdot}4(\lambda/C_s)^{1/4}. \tag{3.3}$$

The resolving power is given by the Abbe expression, since we have seen that the aberration leaves the resolving power unchanged:

$$d = 0{\cdot}6\lambda/\alpha$$

or putting α equal to the above value:

$$d = 0{\cdot}43C_s^{1/4}\lambda^{3/4}. \tag{3.4}$$

Using the spherical aberration data described in Chapter I, this optimum resolving power can be computed. In Figure 3.3 the resolving power given by equation (3.4) is

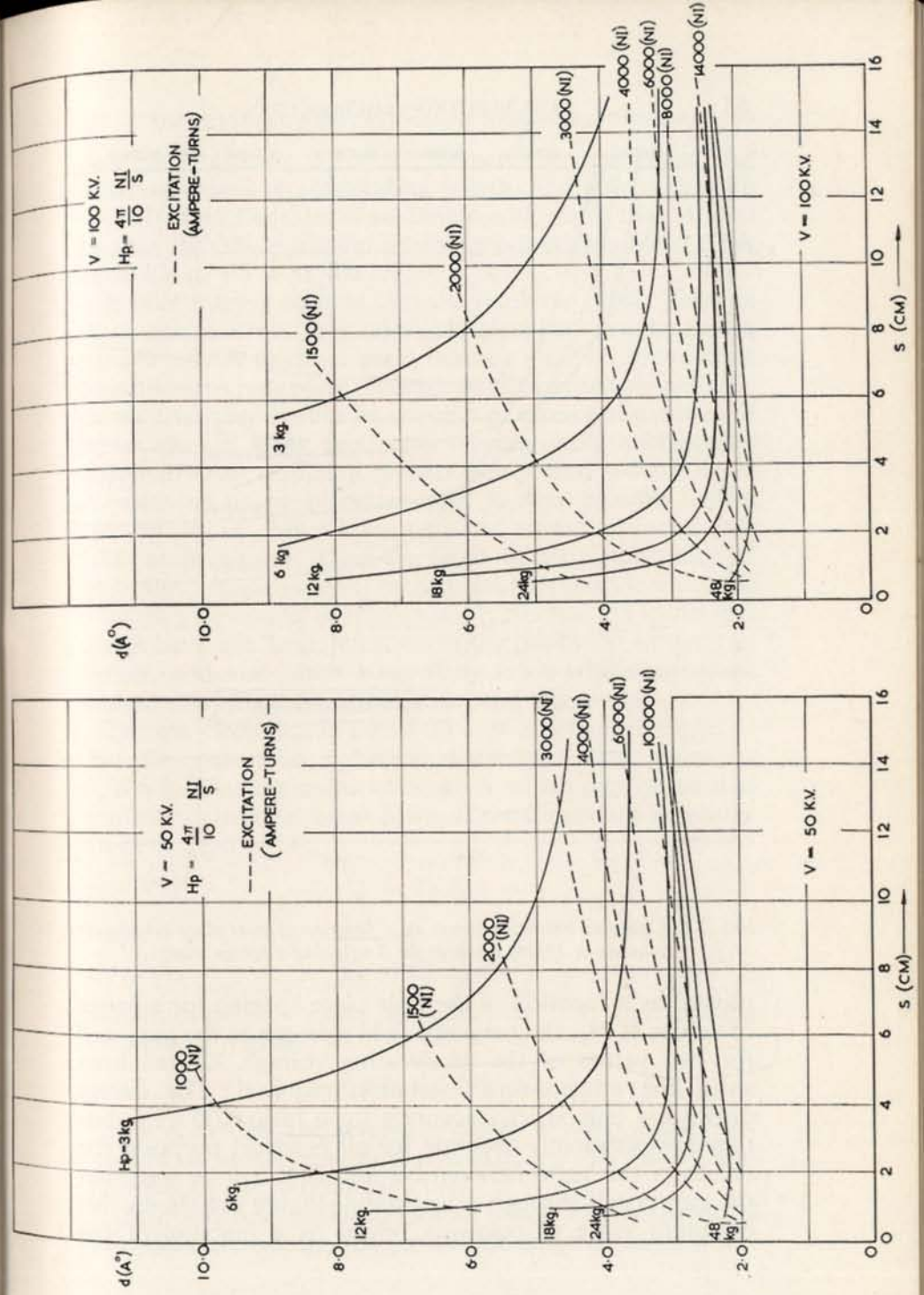


FIG. 3.3. *Resolving power as a function of pole piece spacing, magnetic field and excitation for 50 keV and 100 keV electron energies.*

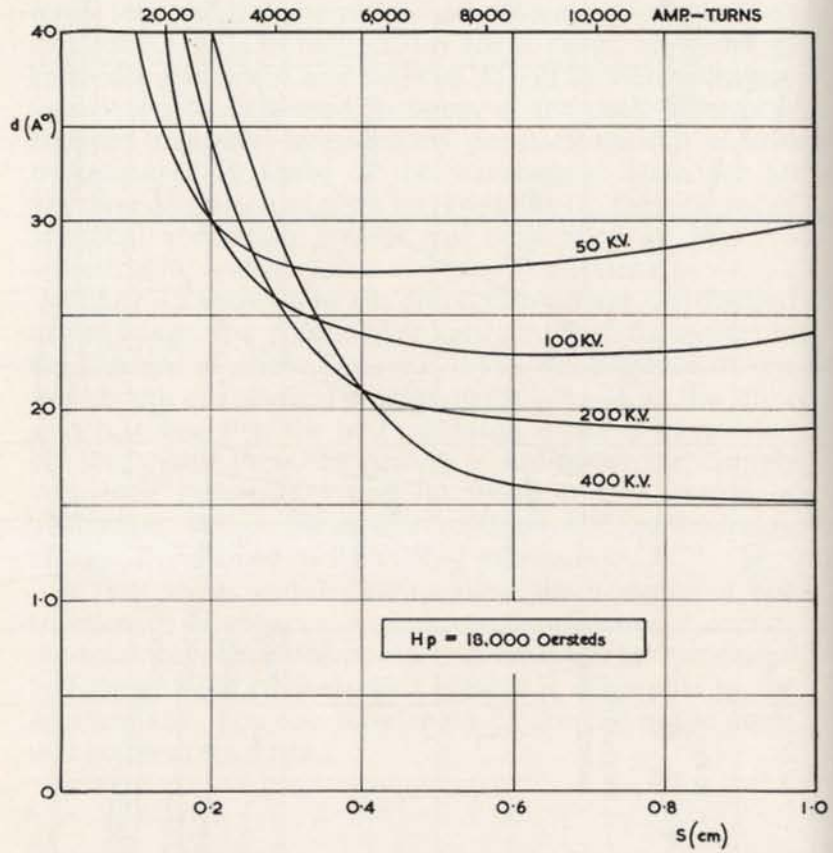


FIG. 3.4. *Optimum resolving power as a function of pole piece spacing or excitation at 18,000 oersteds for 3 values of electron energy.*

plotted as a function of the pole piece spacing for a range of values of H_p, the magnetic field strength in the gap, and for two values of the accelerating voltage. Dotted lines show the ampere-turn excitation required. The curves drawn are only strictly accurate for a ratio $S/D = 1{\cdot}0$, but they are sufficiently accurate for all practical purposes for values of this ratio between 0·2 and 2. It will be seen that the curves fall through a minimum. Figure 3.4 shows this optimum value of resolving power as a function of the

pole piece spacing S, or the ampere-turn excitation, for different values of accelerating voltage and a field strength of 18,000 oersteds, as is realisable with good iron. It will be seen that the optimum resolving power falls from 2·7 Å at 50 kV to 1·5 Å at 400 kV.

It is of interest to note that the resolving power given in these curves is not very critically dependent on the design parameters. A study of most existing electron microscope objective lens pole piece dimensions and excitation facilities, shows that they should be capable of a resolving power of better than 5 Å. In fact, only recently have microscopes appeared with resolving powers better than about 30 Å. A resolving power of better than 10 Å is possible with a flux density of only 4 kilogauss, and an excitation of only 1000 ampere-turns. Over-emphasis of the importance of high flux densities and inadequate appreciation of the principles of design of magnetic circuits has led to designs which have lost large fractions of the excitation in the iron circuit, and also given large stray fields with consequent undesirable effects on astigmatism and alignment.

For the electrostatic Einzel lens, it was seen (Chapter I) that the minimum spherical aberration was given by $C_s/S \approx 5$. The minimum value of S which can be used is limited by electrical breakdown. If the maximum allowable stress is given by X_m

$$S = V/X_m.$$

Then

$$C_s = 5V/X_m.$$

Substituting this value in equation 3.4 and putting $\lambda = \sqrt{(150/V_r)}$ gives:

$$d = 4{\cdot}2/X_m^{1/4}V_r^{1/8}.$$

It is seen that the resolving power is only slightly dependent on voltage and field strength; for $V_r = 100$ kV and $X_m = 200$ kV/cm, d becomes 15 Å.

Thus, the best resolving power which can be expected from the electrostatic instrument is nearly ten times worse

than for the magnetic instrument. The difficulties of keeping the electrodes clean and maintaining adequate vacuum to prevent breakdown also count against the use of the electrostatic instrument, though it has the important advantage of requiring no special high-voltage stability such as is required for the magnetic instrument, since, if the lens is operated from the same potential as the electron gun, the focal length is substantially independent of the variations in that potential.

Resolving Power as Limited by Chromatic Effects

It has already been seen (Chapter I) that the focal length of the electron lens is a function of both its excitation and the electron energy or wavelength. Variation of the focus beyond a certain limit during an exposure will clearly mar the image resolution. On purely geometric considerations deterioration of the image results immediately the focus departs from the optimum. It becomes apparent, however, when the image formation is considered on a wave optical basis, that up to a certain limit, change of focus has only a very small effect on resolution. This is demonstrated in Figure 3.2. It has usually been assumed that the permissible focus variation is that which would introduce a marginal path difference of $\pm\frac{1}{4}$ wavelength, or:

$$\begin{aligned} \tfrac{1}{2}\delta f\alpha^2 &= \pm\lambda/4 \\ \delta f &= \lambda/\alpha^2 \end{aligned}$$

or putting $\alpha = 0{\cdot}61\lambda/d$

$$\delta f = 2{\cdot}7d^2/\lambda.$$

It is true that this change of focus has little effect on resolution, but it can have considerable effect on contrast. This is particularly so, when the phase contrast effects discussed in Chapter IV are considered, since they give an enhancement of contrast slightly off focus which, however, reverses in sign either side of the focus. Thus, if the focus is 'swinging' through the central value due to changing high voltage or lens excitation, the enhancement effect tends to cancel out. To counter this effect, the limit of variation

should be made at least four times smaller than the above, giving:

$$\delta f < 0{\cdot}7d^2/\lambda.$$

In the magnetic lens the variation of focal distance is given by:

$$\delta f = C_c(\delta V/V - 2\delta I/I)$$

where $\delta V/V$ and $\delta I/I$ are the fractional changes in electron gun voltage and lens excitation current respectively and C_c is the chromatic constant. Thus the required limits to these variations are:

$$\delta V/V - 2\delta I/I < 0{\cdot}7d^2/\lambda C_c.$$

Since it is most likely that the variations of voltage and current will be quite unrelated, it is reasonable to allow each of them to rise to a value of 0·7 of the total. Thus, the individual requirements become:

$$\begin{aligned} \delta V/V &< 0{\cdot}5d^2/\lambda C_c \\ \delta I/I &< 0{\cdot}25d^2/\lambda C_c. \end{aligned}$$

It will be seen from Figure 1.10 that for a lens near the minimum focal length, the value of C_c/f is about 0·75. The value of focal length used in an objective lens depends to a large extent on questions of mechanical design (see Chapter VIII). Table 3.1 shows typical values of the limiting high voltage stability in parts per million (p.p.m.) for 80 kV operation both as a product with the chromatic constant (C_c), and with a typical value of $C_c = 0{\cdot}4$ cm.

TABLE 3.1

High-voltage stability requirements as a function of resolving power

d (Å)	2	4	8	16
$C_c(\delta V/V)$ (C_c in cm; $\delta V/V$ in p.p.m.)	0·48	1·9	7·8	31
$\delta V/V$ in p.p.m. ($C_c = 0{\cdot}40$ cm)	1·2	4·8	19	78

It is seen that the requirement is extremely stringent. The requirement for current stability is, of course, twice as severe. It is now common practice to talk of commercial microscopes with a resolving power approaching 10 Å requiring a voltage stability better than about 30 p.p.m. and current stability better than 15 p.p.m. for optimum contrast performance. Such a stability requires the most exacting design of the stabilising equipment. To obtain a resolving power of 5 Å, and to obtain good contrast, requires a four times improvement over these figures and may require resort to special techniques (Haine and Jervis, 1955).

The practical aspects involved in obtaining and measuring resolving power are dealt with in Chapter VIII.

Possibilities for Overcoming the Limitations to Resolving Power

It has been shown that the resolving power of the electron microscope, using conventional lenses, is limited to a value of about 2 Å. Some doubt may be expressed as to whether any improvement on this is worth while since it is already in the region of atomic dimensions, and practical limitations (see Chapter V) have already reached formidable proportions. However, present appearances are often shown deceptive in the light of future advance. It is likely that the main gain which might be made by a further reduction in spherical aberration would be an improvement in contrast. The use of a larger beam angle would, of course, give a brighter image but only at the expense of still higher loading of the object.

At the present time the achievement of spherical aberration correction is more likely to prove useful in allowing greater intensities in electron probe instruments, such as the X-ray microscope (Cosslett, 1957), and the micro-probe X-ray analyser (Castaing, 1954), than in providing improved electron microscopes.

There have been a number of methods described for the correction of spherical aberration. Gabor (1944) has

suggested a method making use of space charge trapped in the axial region of the lens and also a method using an axial electrode (1946). Neither of these appear to offer very great hope of practical realisation. Scherzer (1949) has pointed out that the inclusion of thin foils, transparent to electrons, which allow the electrodes to be continuous through the axis, theoretically allow the design of an aberration-free lens. An application of this suggestion has been discussed by Gianola (1950). The difficulty of providing foils sufficiently thin to give negligible electron scattering and yet strong enough to withstand the electrostatic forces makes this system appear not very promising. Kompfner (1941) has suggested a method based on the use of a variable transit time produced in a pulsed electron beam with a lens excited with a time varying potential. Modifications of this method have been suggested by Scherzer (1947) and by Zworykin *et al.* (1945). Gabor (1944) has pointed out that the requirement for maintaining coherence renders the systems suggested useless for microscopy. Scherzer has shown that this objection is not necessarily inherent in the method. The difficulties are very great. The excitation frequencies required would be in the thousands of megacycles range.

A more promising method proposed by Scherzer (1949 a) uses lens elements of non-rotational symmetry. Basically the proposed system comprises a compound lens made up of elements of non-axial symmetry. Two elements are involved, both having electrodes arranged around the axis alternately positive and negative in potential. The first element has four such electrodes and the second eight. The first type of element acts as a cylindrical lens, while the second can be used to provide negative spherical aberration. A further modification of the original system makes use of a combination of normal symmetrical lenses, Scherzer's eight-pole correctors and cylindrical Einzel lenses. Figure 3.5 shows such a combination used in experiments by Möllenstedt (1954). Möllenstedt's experiments have shown that such a system can provide a lens of lower aberration than conventional lenses, but it still remains to be shown

that such a complex system can be operated with sufficient stability for high resolution electron microscopy. Further analyses of this type of system have been described by Archard (1954) and by Burfoot (1954).

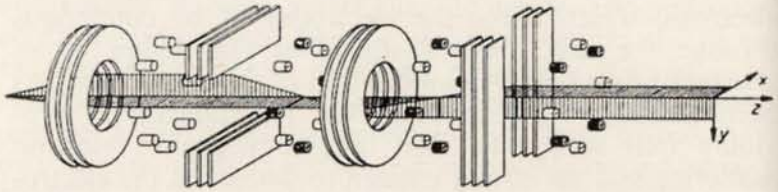
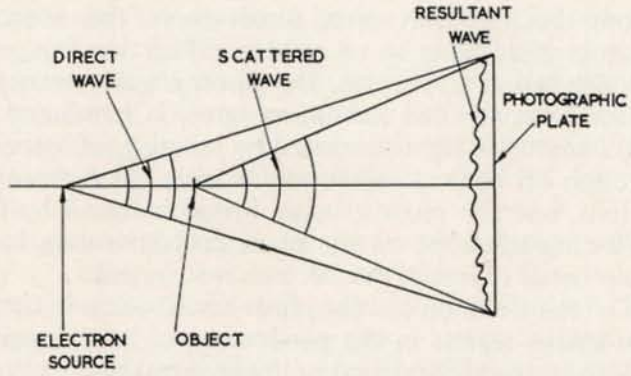


FIG. 3.5. *Schematic diagram of system for the correction of spherical aberration (Möllenstedt, 1954).*

Another method for compensating the lens aberration has been proposed by Gabor (1949). This method constitutes, what amounts to, a new method of image formation which was called by Gabor 'Microscopy by Reconstructed Wave Fronts'. Later the method was called 'Diffraction Microscopy'. Gabor pointed out that the information concerning the distribution of scattering power of the object is contained in the scattered wave alone, provided the specimen is illuminated with an adequately coherent beam. Thus, if the distribution of amplitude and phase in the scattered wave can be recorded, the information concerning the distribution of scattering atoms in the object can, in principle, be reconstructed. Since it is not possible to record the phase of the wave in the object, Gabor resorts to an artifice; the scattered wave is added to a much stronger background wave and the amplitude distribution of the resultant wave is recorded on a photographic plate. This distribution, he shows, contains the required information to an adequate degree of accuracy. The distribution is in the form of a pattern of diffraction fringes following the general contours of the object, the fringes are distorted by the presence of the lens aberrations. In order to reconstruct a focused image of the object the recorded pattern, 'the diffraction image', in the form of a transparency, is irradiated with a coherent beam of light whose



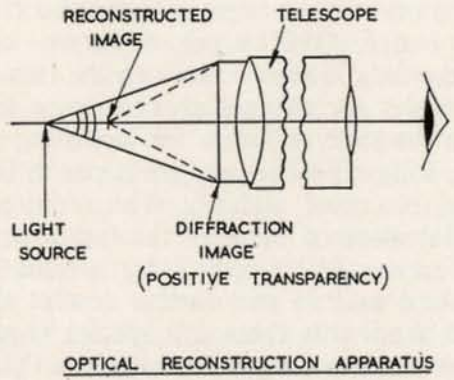


FIG. 3.6. *Schematic diagram illustrating the principle of the Gabor diffraction microscope system.*

wave front is distorted by aberrations in the optical system equivalent to those in the electron optical system; a telescope focused on a certain plane in this system will observe a reconstructed focused image of the object.

The system as originally proposed by Gabor is illustrated in Figure 3.6. Here the object is irradiated from a minute electron source which produces a coherent wave, part of which is scattered by the object and interacts with the

strong background wave, direct from the source, at a photographic plate to record the diffraction image. In the reconstruction apparatus, the photographic transparency, which Gabor called the 'hologram', is irradiated from a small source of light produced by imaging an actual source through an optical system containing the necessary aberrations, and the reconstructed image is found by focusing a viewing telescope on the plane corresponding to that of the original object in the electron instrument.

The failure to record the phase distribution in the diffraction image results in the production of two reconstructed images in the reconstruction apparatus, these images being displaced in axial position by an amount depending on the distance between the original electron source and the object. One of these images is free from aberration while the other contains twice the aberration of the uncorrected image. In the plane of the corrected image appears also the out-of-focus second image. This produces some confusion, particularly if the axial distance between the two images is small. The necessity for using high coherence leads to a low intensity in the electron beam for recording purposes, and therefore to long exposures and the necessity for greater mechanical and electrical stability. The requirement for keeping the axial distance between the two images in the reconstruction apparatus large, still further increases the coherence required and so still further lowers the image intensity. Apart from this there are special requirements for the specimen which must not occupy more than a small fraction of the total field, lest it should unduly interfere with the background wave which is used for reference. A corollary to this is that the specimen must not be mounted on a highly scattering support film since this would destroy the coherence of the background wave.

A further limitation to the method, as proposed by Gabor, has been pointed out by Haine and Mulvey (1952). The Gabor method requires illumination of the object from an electron source whose Gaussian size* must be

* Gaussian size ≡ the size which would be obtained in the absence of aberrations.

comparable with the resolving power required. In order to produce such a small source, a demagnifying electron optical system is required, comparable with the electron optical system of the conventional electron microscope. The use of this system sets the same condition for high voltage and lens current stability as required for the conventional instrument, except that for points in the image which are off the axis an even more stringent voltage stability is required, because of an effect which may be likened to the chromatic difference in magnification. This effect results in a very severe limitation to the useful image field and renders the method extremely difficult in practical application. Haine and Dyson (1950) have suggested a modification to the Gabor method which eliminates the above difficulty. In the modified method (the transmission method), an electron microscope is operated in the conventional manner, except that the illuminating beam is made very coherent (very near parallel) and the image is recorded out of focus, so that the Fresnel diffraction pattern modified by the lens aberration is recorded. Gabor's theory of reconstruction can just as well be applied to this diffraction image, and modified methods of optical reconstruction, including the introduction of the necessary compensating aberration, have been proposed by Dyson (1950) and Gabor (1951). Haine and Mulvey (1952) have shown that the chromatic effects present in the projection method are negligible in the transmission method and that the stability conditions for a given resolving power are identical to those for the normal transmission electron microscope. A further advantage of this system is that the whole instrument can be set up and aligned as a conventional instrument, before applying it to the diffraction method.

The necessary conditions for illuminating the object can be obtained by the use of a very small source, but again the necessity for spacing the double reconstructed images some distance apart requires a very high coherence which leads to low illumination intensity, and exposure times of the order of minutes.

Experimental attempts by Haine and Mulvey (1954) have

shown that reconstructed images can be obtained, but a good deal more experimental work would be necessary to prove any real practical advantage over the conventional method.

REFERENCES

ARCHARD, G. A. *Proc. Internat. Conf. on Electron Microscopy*, London (1954).
BURFOOT, J. C. *Proc. Internat. Conf. on Electron Microscopy*, London (1954).
CASTAING, R. *Proc. Internat. Conf. on Electron Microscopy*, London (1954). Published by The Royal Microscopical Society.
CONRADY, A. E. *Monthly Notices Roy. Astronom. Soc.*, 575, June, 1919.
COSSLETT, V. E., ENGSTRÖM, A. and PATTEE, H. H. *X-Ray Microscopy and Microradiography*, Academic Press (1957).
DYSON, J. *Compt.rend. 1-ier congr. intern. microscopie electronique*, Paris (1950).
GABOR, D. *The Electron Microscope*, Hulton Press, London (1944).
GABOR, D. *Nature*, **158** (1946), 198.
GABOR, D. *Proc. Roy. Soc.* (London), A.**197** (1949), 454.
GABOR, D. *Proc. Phys. Soc.* (London), **64** (1951), 449.
GIANOLA, U. F. *Proc. Phys. Soc.* (London), **63** (1950), 1037.
HAINE, M. E. and DYSON, J. *Nature* (London), **166** (1950), 315.
HAINE, M. E. and JERVIS, M. W. *Proc. I.E.E.*, B.**102** (1955), 265.
HAINE, M. E. and MULVEY, T. *J. Optical Soc. Amer.*, **42** (1952), 763.
HAINE, M. E. and MULVEY, T. *J. Sci. Instrum.*, **31** (1954), 326.
KOMPFNER, R. *Phil. Mag.*, **32** (1941), 410.
MÖLLENSTEDT, G. *Proc. Internat. Conf. on Electron Microscopy*, London (1954).
SCHERZER, O. *Optik*, **2** (1947), 114.
SCHERZER, O. *J. Appl. Phys.*, **20** (1949), 1.
SCHERZER, O. *Optik*, **5** (1949 a), 497.
ZWORYKIN *et al. Electron Optics and the Electron Microscope*, John Wiley, New York (1945).

CHAPTER IV

IMAGE CONTRAST

IN the optical microscope, contrast arises mainly from differential absorption of light between one part of the object and another. In the electron microscope, any appreciable amount of absorption is accompanied by a large temperature rise and considerable scatter in energy of the transmitted electrons; the latter leading to loss of resolution due to chromatic effects. It is thus necessary to make the electron microscope specimen very thin to minimise absorption. Contrast must then arise from some mechanism other than absorption.

For many years it was thought that the most important mechanism was a loss of electrons from the image in those parts of the object where angular scatter was sufficient to deflect electrons outside the aperture of the objective lens. Thus a highly scattering part of the object would appear dark. It is now known that the mechanism of contrast is much more complicated than this. Contrast arises from a combination of several effects associated with the electron scattering and the optical properties of the microscope system.

The scattering of electrons by single atoms must provide the basis for any calculation, at least where elastic scatter is concerned. For inelastic scatter the interatomic fields and collective effects, due to the conduction electrons, may be of great importance. Most published work has dealt with single atom objects, though the way is now open for a more precise extension to multiple atom objects. Earlier work which should be mentioned, mainly as historical interest, is that due to Marton (1934 and 1936), von Ardenne (1938), and Hillier (1941), who applied the classical scattering theory to determine the probability of electrons being scattered outside the microscope aperture. A later paper

by Marton with Schiff (1941) described the application of elastic and inelastic scattering to the contrast problem. Boersch (1946 and 1947), and also von Borries (1949), took the argument a good deal further, but presented theories still based on a limited consideration of all the factors involved. Glaser (1952) has described the general wave optical considerations involved in image formation, but has made no detailed application. Hall (1953) describes the derivation of contrast values from total atomic cross-sections for extended objects, but neglects the wave optical effects. The present author (1957) has made some attempt to consider the overall problem on a wave optical basis with a simplified theory which leads to easy evaluation for single atom objects and could be easily extended to more complex objects. This theory should be adequate for elastic scattering effects but, as will be seen, needs extension to deal adequately with inelastic scattering.

Electron Scattering

Looked at from a classical standpoint an electron in passing through the electrostatic field of an atom can be deflected through an angle which increases as the radial distance from the nucleus to the electron path decreases. On being deflected the electron momentum changes. The momentum of the atom must also change to conserve total momentum and, therefore, some energy transfer from electron to atom must take place. We can describe the angular scatter and energy transfer as two effects and, because the transfer is small, the scatter of this type is termed elastic or coherent since the energy loss is not, in general, enough to affect the wave packet coherence.

For radial distances between electron path and nucleus comparable with the outer electron levels the interaction is mainly with the electronic field. The scatter angle is small but, since the scattering mass of electrons is small, there is a relatively higher probability of energy loss. This process is termed inelastic scattering as the scattered electrons may have lost so much energy as to be no longer coherent with the original beam.

It will be appreciated that there is no sharp dividing line between these various processes, but in the present state of our understanding of the overall process it is convenient to divide it up into a number of separate processes, to access the results of each, and then to combine these results to obtain the total effect.

To carry out this evaluation we require to know the angular distribution of elastically scattered electrons, the angular and velocity distribution of the inelastically scattered electrons, and the energy transfer to the scattering atoms. Unfortunately, we still have insufficient data to enable a very complete analysis to be made in the case of inelastic scattering.

Elastic Scattering

The original calculation by Rutherford on classical lines has been considerably modified by the application of wave mechanical methods. Large angle scattering still agrees with the classical results, but at small angles the wave mechanical methods show large deviations (e.g. Mott and Massey, 1933). For the very small angles of interest in the electron microscope, Lenz (1954) has shown even the earlier wave mechanical calculations to be in error. He derives an expression for the angular distribution of scattered electrons from a beam interacting with a single atom:

$$I_\theta = \frac{4}{a_H{}^2} \frac{Z^2}{(q^2 + 6Z/\Theta)^2} \tag{4.1}$$

where I_θ is the fraction of the electrons falling on one square centimetre in the illuminating beam which will be found in a hollow cone of unit solid angle making an angle θ with the direction of the beam, a_H is the Bohr radius of the hydrogen atom ($=0{\cdot}529$ Å), Z is the atomic number,

$$q = (4\pi/\lambda) \sin \theta/_2 \quad \text{and} \quad \Theta = \int_0^\infty \rho_r r^2 . 4\pi r^2 \, dr$$

where ρ_r is the charge distribution in the atom.

The value of $6Z/\Theta$ can be derived from Lenz's data as:

$$6Z/\Theta = Z\lambda^2\phi_0/1{\cdot}4\times10^{-14}\pi a_H{}^2 \qquad (4.2)$$

where $e\phi_0$ is the electron energy in electron volts.

Putting

$$\lambda = \sqrt{150/\phi_0}\times10^{-8}\text{ cm}$$
$$6Z/\Theta = 0{\cdot}34Z/a_H{}^2\text{ cm}^2. \qquad (4.3)$$

One interesting result from Lenz's theory is that at very small angles where $q^2 \ll 6Z/\Theta$ the scattering cross-section becomes independent of angle, atomic number and energy, and is equal to about 10^{-15} cm^2.

At large angles where $q^2 \gg 6Z/\Theta$ equation (4.1) gives:

$$I_\theta = 12{\cdot}6\times10^{-16}Z^2/\phi_0{}^2\sin^4\theta/2. \qquad (4.4)$$

This is the classical Rutherford scattering formula.

The limit where the angular dependent term equals the constant term in the denominator of (4.1) is given by:

$$\theta \approx 2{\cdot}1\sqrt{Z/\phi_0}$$

or for $Z=8$ and $\phi_0=10^5$ volts, $\theta=0{\cdot}02$ rad; and for $Z=64$ and $\phi_0=10^5$ volts, $\theta=0{\cdot}05$ rad.

It is seen, therefore, that the intensity falls little over the aperture angle used in the electron microscope (0·002–0·005 rad).

The total cross-section for elastic scattering, or the fraction of all electrons which fall on 1 cm^2 of image surface which are scattered at all, is given from Lenz's data as:

$$\underline{\sigma_e = 8{\cdot}5\times10^{-15}\,A_t/\phi_0} \quad \text{(non-relativistic)}$$

where A_t = Atomic weight.

A further parameter of interest is the thickness of specimen film which gives a probability of scatter of $1/e$. This is usually measured in grammes/cm^2 and is given by:

$$\gamma' = (A_t/\sigma_e)\times\text{proton mass}$$
$$\underline{= 2\times10^{-10}\phi_0.}$$

The actual thickness is then given by:

$$t_0 = 2\times10^{-10}\phi_0/D$$

where D is the density.

It is seen that the density thickness varies from about 1000 Å for carbon to less than 100 Å for heavy elements.

N.B.—There is a slight approximation involved in the derivation of σ_e and γ' above, but the errors are not of practical significance.

Inelastic Scattering

The data on inelastic scattering is nothing like so complete as for elastic. Again Lenz's data is probably most reliable, but he has only derived the angular distribution of inelastically scattered electrons and not the corresponding energy loss distribution. The expression for the angular distribution given by Lenz is:

$$I_\theta' = \frac{4Z}{a_H{}^2 q^4}\left[1-\frac{1}{(1+q^2\Theta/6Z)^2}\right]. \qquad (4.5)$$

where q now has a value:

$$q = (2\pi/\lambda)\sqrt{\theta^2+(J/4e\phi_0)^2} \qquad (4.6)$$

where J is the ionisation energy of the atom.

For parameter values of interest $q^2\Theta/6Z \ll 1$ and (4.5) becomes:

$$I_\theta' = \frac{4Z}{a_H{}^2q^4}\times2q^2\Theta/6Z$$
$$= 4\Theta/3a_H{}^2q^2$$

putting in the value of Θ from (4.2) and the value of q from (4.6) gives:

$$I_\theta' = 8{\cdot}8\times10^{-15}\phi_0/(6{\cdot}3+\phi_0{}^2\theta^2)\text{ cm}^2$$

where J is given a typical value of 10 electron volts.

In order to account for the energy spread of the inelastically scattered atoms it can be assumed that they will have a Gaussian distribution with a mean width $e\phi_m$.

Thus the total distribution in angle and energy is then:

$$\mathscr{I}_{\text{in}}\,\mathrm{d}\phi_{\text{e}} = \frac{8{\cdot}8\times10^{-15}}{6{\cdot}3+\phi_0{}^2\theta^2}\exp\left(\frac{\phi_{\text{e}}}{\phi_{\text{m}}}\right)^2\mathrm{d}\phi_{\text{e}} \qquad (4.8)$$

where $e\phi_0$ is the loss in energy.

The total cross-section for inelastic scatter is given by Lenz as:

$$\sigma_{\text{in}} = \frac{\lambda^2\Theta}{3\pi a_{\text{H}}{}^2}\log_{\text{e}}(6Z/\bar{q}^2\Theta)$$

where $\bar{q}$ is given by the smallest possible value of q:

$$\bar{q} = \pi J/2\lambda\phi_0$$

putting in this value and that of Θ from (4.2):

$$\underline{\sigma_{\text{in}} = 2{\cdot}8\times10^{-14}\phi_0{}^{-1}\log_{\text{e}}(73Z\phi_0/J^2).}$$

Lenz's calculations for inelastic scatter have received reasonable experimental confirmation (Biberman, 1949; Leonhard, 1954; Haine and Agar, 1959).

We now proceed to use the expressions for elastic and inelastic scatter to derive the expected contrast for single atom objects.

Contrast

Before proceeding, a brief consideration should be given to the historical method for calculating contrast. It has been argued that the contrast must be produced by the electrons lost to the aperture. The fraction lost is given simply by:

$$\int_{\theta_0}^{\infty}2\pi\mathscr{I}_\theta\sin\theta\,\mathrm{d}\theta = \sigma_{\text{e}} - \int_0^{\theta_0}2\pi\mathscr{I}_\theta\sin\theta\,\mathrm{d}\theta.$$

To calculate the resulting contrast it is assumed that the electrons are all missing from the resolution area around the Gaussian image of the scattering atom. Thus, the contrast is given by:

$$C = \left[\sigma_{\text{e}} - \int_0^{\theta_0}2\pi\mathscr{I}_\theta\sin\theta\,\mathrm{d}\theta\right]\Big/d^2.$$

The contrast value given by this expression is substantially correct, but the argument leading to it is only of limiting validity as it disregards the phase effects which may, and do, have profound influence on contrast. In fact, the contrast given by the above expression is only correct on focus; a very slight change of focus will result in a large contrast change.

Before attempting to calculate contrast effects it will be helpful to add a few remarks to explain how the scattering effects just described can be interpreted in terms of the optical theory where it is more usual to consider refractive index variations, producing a variation of optical path and absorption.

It has already been explained in Chapter II how a small phase object with no absorbing power but of refractive index or optical path length different from its surroundings produces a complex wave which can be resolved into a plane wave and a scattered wave. The amplitude distribution of the scattered wave was related to the variation of optical path length in the object and the propagation of this wave calculated by a double Fourier transform. The atom could well be described in terms of a certain distribution of optical path length varying with radius from the nucleus, instead of in the terms of angular scatter cross-section, as given by the theoreticians who calculated the data. The angular distribution of scatter is simply the distribution found in the back focal plane of the objective lens, so that, in effect, the calculation for angular scatter distribution already contains the first Fourier transform, and contrast calculations would appear to need only the completion of the second transform with limits to the integral corresponding to the lens aperture, if necessary including path difference terms for aberrations and/or departures from Gaussian focus. However, one difficulty arises in that the scattering calculations which give the angular distribution of intensity of scattered electrons, or the square of the amplitude distribution, do not give the relative phase angle between the scattered wave and the background wave. It will later be seen that this difficulty

becomes of importance only in calculating the in phase or amplitude contrast and that the phase angle may be derived from total energy considerations.

The Coherent Wave

In the image plane the amplitude distribution of the coherent scattered wave produced by elastic or nuclear scattering is given by putting $\psi' = \mathscr{I}_\theta^{1/2}/f$ in equation (2.6a):

$$\psi'' = \int_0^{r_0'} \frac{2\pi \mathscr{I}_0^{1/2} r'}{f^2\lambda} J_0\left(\frac{2\pi r' r''}{f\lambda}\right) \mathrm{d}r'$$

where r_0' is the radius of the objective aperture and aberrations and defocusing are neglected.

It has been seen that the value of $\mathscr{I}_0$ is substantially constant and equal to 10^{-15} cm^2 across the aperture angles relevant to the electron microscope. The integral then gives:

$$\psi'' = \frac{\mathscr{I}_\theta^{1/2} r_0'}{r'' f} J_1\left(\frac{2\pi r'' r_0'}{f\lambda}\right). \tag{4.9}$$

This expression gives the amplitude distribution of the scattered wave, or the part of it transmitted by the objective aperture, in the image plane. To obtain the image contrast the scattered wave must be compounded with the background wave, but this is not possible until the relative phase between the two waves is known.

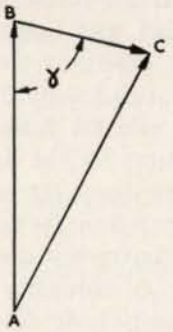


FIG. 4.1. *Vectorial representation of background wave* $\vec{AB}$, *total wave* $\vec{AC}$ *and scattered wave* $\vec{BC}$.

The Phase Angle γ

We can now proceed to evaluate the phase angle γ by energy considerations. The energy loss to the objective aperture must be equal to that missing from the image. The former is given by:

$$\delta\mathscr{E}_a = \int_{\theta_0}^{\infty} 2\pi \mathscr{I}_\theta \sin\theta \, \mathrm{d}\theta$$

or

$$\delta\mathscr{E}_a = \sigma_e - \int_0^{\theta_0} 2\pi \mathscr{I}_\theta \theta \, \mathrm{d}\theta$$

$$\delta\mathscr{E}_a = 1{\cdot}36 \times 10^{-14} Z/\phi_0 - \pi \mathscr{I}_{\theta_0} \theta_0^2 \tag{4.10}$$

The energy in the image is given by integrating the square of the image amplitude.

The vectorial representation of the wave amplitude is shown in Figure 4.1. $\vec{AB}$ is the unity amplitude background wave, $\vec{BC}$ is the scattered wave ($\angle ABC = \gamma$), and $\vec{AC}$ is the resultant wave amplitude.

We have:

$$\psi_t^2 = 1 - 2\psi'' \cos\gamma + \psi''^2. \tag{4.11}$$

The energy absent in the image is then:

$$\delta\mathscr{E}_i = \int_0^{\infty} 4\pi r'' \psi'' \cos\gamma \, \mathrm{d}r'' - \int_0^{\infty} 2\pi r'' \psi''^2 \, \mathrm{d}r''.$$

Putting in the value of ψ'' from equation (4.9):

$$\delta\mathscr{E}_i = \int_0^{\infty} \frac{4\pi \mathscr{I}_{\theta_0}^{1/2} r_0'}{f} \cos\gamma J_1\left(\frac{2\pi r'' r_0'}{f\lambda}\right) \mathrm{d}r''$$

$$- \int_0^{\infty} \frac{2\pi \mathscr{I}_\theta r_0'^2}{r'' f^2} J_1^2\left(\frac{2\pi r'' r_0'}{f\lambda}\right) \mathrm{d}r''.$$

Remembering

$$\int_0^{x_0} J_1(x) = 1 - J_0(x_0)$$

and

$$\int_0^\infty \frac{J_1^2(x)}{x} = \tfrac{1}{2}[J_0^2(x)+J_1^2(x)]$$

and

$$J_0(0) = 1, \quad J_1(0) = 0, \quad J_1(\infty) = J_0(\infty) = 0$$

the integrals give:

$$\delta\mathscr{E}_i = 2\lambda \mathscr{I}_{\theta_0}{}^{1/2} \cos\gamma - \pi \mathscr{I}_\theta \theta_0{}^2. \tag{4.12}$$

Equating (4.11) and (4.12) gives:

$$\cos\gamma = \frac{1{\cdot}36\times10^{-14}Z}{2\lambda \mathscr{I}_{\theta_0}{}^{1/2}\phi_0}$$

putting

$$\mathscr{I}_{\theta_0} = 10^{-15} \quad \text{and} \quad \lambda = \sqrt{\frac{150}{\phi}}\times10^{-8}$$

$$\cos\gamma = 1{\cdot}8Z/\phi_0{}^{1/2} \tag{4.13}$$

e.g. for $\phi_0 = 100$ keV

$$\cos\gamma = 0{\cdot}005Z \tag{4.14}$$

i.e. the scattered wave is almost in quadrature phase relation with the background.

Evaluation of Contrast

Amplitude Contrast

The contrast (C_A) in the image of a single atom is given by the difference between the total and the background intensity divided by the background intensity:

$$C_A = 1 - \psi_t{}^2.$$

Putting $\psi_t{}^2 = 1 - 2\psi'' \cos\gamma + \psi''^2$ (from (4.11))

$$C_A = 2\psi'' \cos\gamma + \psi''^2.$$

The maximum contrast will clearly occur at the centre of the Airy pattern ($r'' = 0$). Putting $r'' = 0$ in equation (4.9)

and substituting in the above with $\mathscr{I}_{\theta_0} = 10^{-15}$ and $\theta = 0{\cdot}61\lambda/d$ gives:

$$C_A = \frac{350}{d^2\phi_0}\left(Z - \frac{23}{d^2}\right) \quad (d \text{ in Å}). \tag{4.15}$$

The following table evaluating (4.15) shows that the amplitude contrast for single atoms is very small ($\phi_0 = 10^5$ volts).

TABLE 4.1

Variation of amplitude contrast with resolving power and atomic number

Z	d^2		
	3	5	7
4	0·11 %	0·09 %	0·06 %
16	1·2	0·45	0·23
64	4·8	1·8	0·9

Phase Contrast

Since the scattered wave traverses the back focal plane of the objective lens over the whole area of the used aperture, whereas the background wave is concentrated mainly very near the axis, the possibility arises for modifying one part relative to the other. In particular the relative phases may be altered and, if desired, the amplitudes. In the optical microscope, methods for controlling both of these parameters have been applied with the result that for phase (non-absorbing) objects very great improvement in contrast has resulted (phase contrast). The reason for this improvement is clear from Figure 4.1. If the angle γ is near 90° the scattered wave amplitude only affects the total wave amplitude by a second order amount. If, however, the scattered wave could be turned in phase by

90° so that it was in phase with the background, its effect would be first order. The contrast would then be:

$$C_{ph} = (1+\psi'')^2 - 1$$
$$= 2\psi''.$$

For $r'' = 0$

$$C_{ph} = 2\pi \mathscr{I}_\theta^{1/2} \theta_0^2 / \lambda.$$

For $\mathscr{I}_\theta = 10^{-15}$, $\phi_0 = 100$ keV

$$C_{ph} = 27{\cdot}6/d^2\% \quad (d \text{ in Å}) \tag{4.16}$$

which is seen to exceed the amplitude contrast by a considerable factor.

Out of Focus Phase Contrast (C_{ph}')

If the microscope objective lens is defocused by a small amount, a phase difference is introduced into the spherical wavefront given by $\beta(\theta/\theta_0)^2$ where β is the phase difference when $\theta = \theta_0$. This phase difference is clearly of most significance near the periphery of the aperture and, therefore, affects the scattered wave much more than the direct wave. It is seen, therefore, to cause a rotation of the scattered wave vector and will, therefore, produce some phase contrast effect. At the same time the image is defocused, but the defocusing effect can be negligible if it is within the depth of focus of the instrument ($\beta \sim \pm\pi$).

The amplitude of the scattered wave in the defocused image in the image plane is given by:

$$\psi''_d = \frac{2\pi}{\lambda} \int_0^{\theta_0} \mathscr{I}_\theta^{1/2} \theta \sin \beta \left(\frac{\theta}{\theta_0}\right)^2 d\theta$$

Since I_θ is substantially uniform over $0 - \theta_0$:

$$\psi''_d = \frac{2\pi \mathscr{I}_\theta^{1/2}}{\lambda} \int_0^{\theta_0} \theta \sin \beta \left(\frac{\theta}{\theta_0}\right)^2 d\theta$$

$$\psi''_d = \frac{\pi \mathscr{I}_\theta^{1/2} \theta_0^2}{\lambda} \frac{1-\cos\beta}{\beta}.$$

The resulting contrast C_{ph}' is given by $2\psi''_d$:

$$C_{ph}' = \frac{2\pi \mathscr{I}_\theta^{1/2} \theta_0^2}{\lambda} \frac{1-\cos\beta}{\beta}$$
$$= C_{ph} \left(\frac{1-\cos\beta}{\beta}\right)$$

Putting $\cos\beta \approx 1 - \frac{1}{2}\beta^2$

$$C_{ph}' = \tfrac{1}{2}\beta C_{ph}.$$

It is seen that a substantial proportion of the full phase contrast is obtained by slight defocusing without loss of resolution. This result is of the greatest significance, indicating that the main source of contrast, at least for fine detail, arises in this manner.

It will also be noted that the out-of-focus phase contrast reverses on either side of true focus, a factor which must be considered in estimating the required voltage and current stability (see Chapter III).

Inelastic Scattering and Contrast

The part played by inelastic scattering in the production of contrast has not yet been adequately evaluated. This is partly because there is still no detailed or accurate knowledge of the distribution of inelastically scattered electrons in angle and velocity. This is true for single atom scattering and even more so for scattering from an atomic lattice where the problem is a good deal more complicated. A full solution must be obtained on a wave mechanical basis where the effect of the inelastic scatter on the wave packet must first be determined, and followed by the calculation of the propagation of the complex wave packet, taking into account the coherence relation between the scattered part and the unperturbed background.

Various attempts have been made to estimate the effect but not always have all the factors been included, and no attempts described so far can be regarded as satisfactory. One approximate way of regarding this problem (Haine, 1957) divides the overall problem into two parts. The

inelastically scattered electrons, having lost energy, are considered as being incoherent with the unscattered wave. In this respect they may be considered as being lost from the main wave, thus reducing its amplitude over the area of scattering. This area of reduced amplitude produces a coherent scattered wave as can be seen by considerations similar to those applied for elastic scattering. The scattered wave will be in antiphase with the uniform background wave and will thus produce a dip in the intensity in the image. The incoherent electrons may now be considered on simple geometric considerations and their focusing calculated taking into account the chromatic aberration of the objective lens, which will result in them being spread over an area greater than the area of the intensity dip produced by the coherently scattered wave. Thus, the inelastically scattered electrons will partly 'fill in' the intensity dip, and hence reduce the contrast so produced.

Difficulties involved in completing this calculation arise in the first place because, to calculate the coherently scattered wave, it is necessary to know the area over which the scattering takes place and this is not specifically known. Secondly, although the angular distribution of inelastic scatter is known, the distribution of energy loss and its relation to the angular distribution does not appear to be known.

TABLE 4.2

Variation of contrast C_{in} and C_{in}' with resolving power and atomic number

Z	d (Å)					
	2		4		8	
	C_{in}	C_{in}'	C_{in}	C_{in}'	C_{in}	C_{in}'
	%	%	%	%	%	%
64	5·6	0·2	1·8	0·06	0·4	0·015
416	7·1	0·2	1·8	0·06	0·4	0·015
64	7·1	0·2	1·8	0·06	0·4	0·015

Taking the scattering area, for a single atom, as equal to the radius of the outer electron shell, and assuming a Gaussian energy loss distribution independent of the scattering angle, Haine (*ibid.*) calculated the two components of contrast with the results shown in the table, C_{in} being the contrast due to the elastically scattered wave, and C_{in}' that due to defocused inelastically scattered electrons.

It is seen that the predominant effect is that due to the coherent wave and the magnitude of the contrast is comparable with the phase contrast produced by elastic scatter.

Extended Objects

The contrast which may be expected from single atoms is possibly of little direct practical interest. On the other hand it must be evaluated if a proper understanding of the overall contrast mechanism is to be obtained. As may be expected, the problem becomes complex when real objects are considered. Space and the relatively backward state of the art prevents any full description being presented and the following is limited to a description of how the theory can be developed to large area objects.

To proceed, it is imagined that the single atom object is extended by adding atoms in close proximity to increase the size. At first the added atoms will be within the resolved area of the instrument. In this case the elastically scattered waves will be substantially co-phasial in the Fourier plane and the amplitude in this plane will increase in proportion to the number of atoms present. The phase contrast will, therefore, increase in this proportion also. When the object size approaches the resolved area, the scattered waves arising at the edge of the aperture are not all in phase and the wave amplitude in the Fourier plane (previously almost constant) begins to fall rapidly towards the aperture periphery. As the object size grows much bigger than the resolved area, the wave amplitude in the Fourier plane drops to zero quite near the axis and follows an oscillation across the radius, approximately following the $J_1(r)/r$ form with succeeding half cycles reversing in phase. At first sight these phase reversals would suggest that phase contrast

must now disappear, or be greatly diminished, but the second Fourier transform introduces corresponding changes in phase as the wave amplitude is integrated across the aperture to give ψ''. Phase contrast is thus partly maintained even across an extended object. The exact effect on 'out-of-focus' phase contrast is in need of evaluation. In any case, as the number of scattering atoms increases, the amplitude of the background wave decreases and phase contrast is thus reduced. When the specimen thickness approaches the transparency thickness the background wave effectively disappears and phase contrast with it.

The latter factor is of the greatest importance when one considers the contrast of small perturbations in an otherwise smooth film. Such a perturbation produces a coherent scattered wave which with a strong background wave will give good contrast. If the background wave is absent the contrast is very greatly reduced. This is the main reason why poor contrast is obtained for fine detail in thick objects. The loss of the coherent background is aggravated by inelastic scattering which at low atomic numbers has a bigger cross-section than elastic.

It will be seen that the contrast to be expected with an extended object is that calculated for a single atom multiplied by the number of atoms in each resolution area. This will remain true up to a density where plural scattering commences, i.e. at about the critical thickness. Beyond this thickness multiple scattering processes complicate the contrast calculation, but in any case the specimen should now be considered too thick for useful observation at high resolution. Logically when such thick objects are being considered with resulting loss in effective resolving power, better results should be achieved by stopping down the objective lens to an even smaller aperture than conventionally used. The critical thickness, quoted earlier in this chapter as $\gamma_e = 2 \times 10^{-10} \phi_0$ gm/cm^2, amounts to an actual specimen thickness of several thousand Å.

The contrast due to inelastic scattering from extended objects tends to zero since the overlapping areas occupied

by the incoherent electrons in the image just fill the dips due to the elastic wave.

The more complicated effects which may result due to the regularity of atomic arrays in a specimen are well illustrated by the beautiful results due to Menter (1958) and others, showing how Moiré patterns may be obtained from superimposed atomic lattices of slightly different atomic spacings. This effect can show up individual dislocations in a thin crystalline film.

One final effect should be mentioned under this chapter heading, though it is not specifically a contrast effect. This is the effect of the electrons on the object. Apart from heating, which can be troublesome (von Borries, 1948), the scattering of the electrons is inevitably accompanied by an energy transfer to the object as has been mentioned before. The ionisation produced by inelastic scattering can result in broken valence bonds in the solid. Subsequently rearrangement of the broken bonds can lead to a change in structure. For example, organic molecules can be broken down to form carbon in some cases, in others, more complex molecules can result.

The energy transfer in elastic collisions is imparted to the nucleus. The exact amount of energy transferred depends on whether the scattering mass can be considered as that of the single atom or whether the interatomic forces share the scattering impulse among groups of atoms. In the former case, considerations of conservation of momentum and energy result in an energy transfer ($\mathscr{E}$) for a scatter through an angle θ given by:

$$\mathscr{E} = h^2\theta^2/2M\lambda^2$$

where M is the mass of the scattering atom.

For 60 keV electrons this gives:

$$\theta_1 = 0{\cdot}18\sqrt{A\delta\phi}$$

where A is the atomic weight.

It can be shown (Haine, 1957) that the probability is high for at least one scattering act, during an exposure, to give an angle sufficient to transfer at least 1 eV to the

atom. 1 eV is typical of the binding energy of a surface atom to a solid. Therefore, one may expect at least all the surface atoms to be evaporated during an exposure. Perhaps this effect is nullified by the carbon contamination which is continually building up on the surfaces.

REFERENCES

VON ARDENNE, M. *Z. Phys.*, **111** (1938), 152.

BIBERMAN, L. M., VTOROV, E. N., KOONER, I. A., SUSHKIN, N. G. and YAVORSKII, B. M. *Dokl. Akad. Nauk SSSR*, **69** (1949), 519.

BOERSCH, H. *Monatsh. Chem.*, **76** (1946), 163.

BOERSCH, H. *Monatsh. Chem.*, **76** (1946), 860.

BOERSCH, H. *Monatsh. Chem.*, **78** (1947), 163.

VON BORRIES, B. *Optik*, **3** (1948), 321.

VON BORRIES, B. *Z. Naturf.*, **49** (1949), 51.

GLASER, B. W. *Grundlagen der Elecktronoptik*, Springer Verlag, Vienna (1952).

HAINE, M. E. *Advances in Electronics*, Vol. 6, Academic Press, New York (1954).

HAINE, M. E. *J. Sci. Instrum.*, **34** (1957), 9.

HAINE, M. E. and AGAR, A. W. *Brit. J. Appl. Phys.*, **10** (1959), 341.

HALL, C. E. *Introduction to Electron Microscopy*, McGraw Hill, New York (1953).

HILLIER, J. *Phys. Rev.*, **60** (1941), 742.

LENZ, F. *Z. Naturf.*, **9a** (1954), 185.

LEONHARD, F. *Z. Naturf.*, **9a** (1954), 727 and 1019.

MARTON, L. *Phys. Rev.*, **45** (1934), 527.

MARTON, L. *Physica*, **3** (1936), 959.

MARTON, L. and SCHIFF, L. I. *J. Appl. Phys.*, **12** (1941), 759.

MENTER, J. W. *Phil. Mag. Suppl.*, **7** (1958), 299.

MOTT, N. F. and MASSEY, H. S. W. *The Theory of Atomic Collisions*, Oxford University Press (1933).

CHAPTER V

PRACTICAL FACTORS LIMITING PERFORMANCE

IN this chapter some of the practical factors limiting performance, particularly resolving power, will be discussed. As in the optical microscope the objective lens is of paramount importance. The theoretical properties of lenses have already been described in Chapter I and a variation of instrumental resolving power with design has been evaluated in Chapter III. The theoretical limits to resolving power were expressed in terms of the pole piece geometry, the strength of the magnetic field between the pole pieces or the excitation applied to the gap. To attain these results the lens must be designed so that adequate excitation is available without overheating, and so that no undue fraction of it is lost in the excitation of the iron circuit. Furthermore, the iron circuit must have sufficiently low reluctance to ensure that the field in the bore of the lens within the iron is sufficiently small to exert no appreciable effect on the electron beam. A further requirement is that the axial symmetry of the lens should be very accurately preserved, particularly in the region of the pole pieces, to avoid astigmatism. The symmetry conditions were also discussed in Chapter I.

Astigmatism

The amount of astigmatism permissible in a microscope of given resolving power is somewhat arbitrary; it not only depends on its effect in spoiling resolution, but on the fact that a small amount of astigmatism can produce artefacts in the image due to its directional nature. Certainly, astigmatism should be reduced to a level where the maximum path difference in the wavefront is below a quarter

wavelength, and preferably an even greater reduction is desirable. The path difference due to a longitudinal astigmatism z_a is given by $\Delta = \frac{1}{2} z_a \alpha^2$, hence the maximum permissible magnitude can be expressed as:

$$\frac{1}{2} z_a \alpha^2 < \lambda/4$$
$$z_a < \lambda/2\alpha^2$$

putting $\alpha = 0{\cdot}6\lambda/d$

$$z_a < 1{\cdot}4d^2/\lambda. \qquad (5.1)$$

TABLE 5.1

Limiting values of longitudinal astigmatism z_a as a function of resolving power and voltage

ϕ_0(kV)	d (Å)			
	2	4	8	16
	microns	microns	microns	microns
50	0·01	0·04	0·16	0·64
100	0·014	0·056	0·22	0·88
200	0·021	0·084	0·33	1·33

From the expression (1.12), it is seen that, for an objective lens of maximum strength, the ratio of the longitudinal astigmatism to the symmetry tolerance (z_a/δ) is about 2, so that the tolerances required are about half the dimensions given in the above table. The best mechanical tolerance which can be obtained even by very careful machining techniques is about one micron, suitable only for a resolving power of about 30 Å at 100 kV accelerating voltage.

Astigmatism can also arise from electrostatic or magnetic stray fields with asymmetry about the axis, particularly if such fields are present in the region of the objective lens. Towards the final image the importance of asymmetric fields in producing astigmatism becomes less, since the beam aperture angle becomes smaller. However, such fields can produce image distortion in extreme cases though

this trouble is rarely experienced. Asymmetric fields can arise from a variety of causes. Inadequate design of the objective lens iron circuit can easily introduce a stray magnetic field which can have an asymmetric component. Such effects are minimised if the flux density in the iron is kept as small as possible and if the iron circuit is kept as symmetrical as possible; it is also advisable to avoid any magnetic gaps or joints in the iron circuit in the region of the pole pieces. Electrostatic fields due to charging of insulating contamination can also give asymmetric fields. Such defects are reduced by keeping the metallic wall surrounding the beam as far from the beam as possible and by suitable apertures and masks to shield any insulating surfaces from the possible incidence of stray scattered electrons or ions produced by the electron beam. Careful cleaning of suspect surfaces to remove insulating films is necessary when trouble of this sort is experienced, but a carefully designed instrument should not need such cleaning more frequently than once every few months, unless it has been badly operated or is being used for the most exacting work (resolving power near the theoretical limit).

A further source of asymmetric electrostatic fields arises from contamination of the objective aperture. This aperture has a diameter given by $\sim f \times \alpha_0$ where f is the objective lens focal length and α_0 the chosen optical aperture angle. For the lens quoted as examples in Chapter VIII the aperture diameter would be 30 microns only, through this the whole imaging beam must pass. Furthermore, this aperture must stop a proportion of the scattered electrons. It is therefore particularly liable to contamination and any contamination is liable to produce astigmatism. For many years it was accepted that platinum or platinum alloy was the most suitable material for such apertures. The apertures are manufactured by drilling, or otherwise producing, suitably sized holes in a disc of the material, about 3 millimetres in diameter. Experiments by Haine and Mulvey (1954) showed that platinum apertures were liable to be contaminated with an insulating film extremely difficult to remove, though workers in Professor Ruska's Laboratory

did develop a somewhat difficult chemical cleaning process which appeared to be at least fairly effective. Haine and Mulvey discovered, however, that apertures made in molybdenum could be cleaned very simply by flashing to about 1300°C on a strip heater in the vacuum of an evaporation plant. This cleaning process needs to be repeated after every few hours' operation where the best performance is required.

It is fortunately possible to correct astigmatism, as was first pointed out by Hillier and Ramberg (1947). There are two basic methods of achieving correction. In the first, the lens field is distorted by some means to correct the asymmetry. Thus, Hillier and Ramberg achieved correction by eight adjustable magnetic screws in the gap between their pole pieces. A more satisfactory method now widely applied is by the addition of a weak cylindrical lens of strength just sufficient to correct the cylindrical component of the objective lens, oriented in a perpendicular direction to that of the astigmatism. Leisegang (1953) has achieved this by placing a magnetic component of two-fold symmetry just below the bottom lens pole piece, where the penetration of the lens field into the pole piece bore on to the corrector provides enough lens action to correct the astigmatism; the amplitude of the correction is then adjusted by raising or lowering the corrector, while its orientation is adjusted by rotation.

Bertein (1948) has considered various alternative methods for correction. The most useful is possibly a four-pole electrostatic system illustrated in Figure 5.1. Four short rod electrodes are mounted symmetrically around the optical axis and held at alternate positive and negative potentials. The resulting field has two-fold symmetry, and acts as a cylindrical lens. The advantage of such a system over a simple three-slit electrostatic lens, for example, lies in the fact that the lens action for a given applied potential is very much stronger since it results from the total field amplitude and not just from the fringing component of it. The orientation of the correction is varied by rotation of the whole unit. By using a six-pole system (Figure 5.2),

Rang (1948) was able to obtain the variation of orientation without rotation. In effect, this system comprises two correctors oriented at 45°, and, by adjusting their relative amplitudes, any resultant amplitude and orientation can be obtained. In practice, suitable circuits to give independent control of amplitude and orientation involve the use of

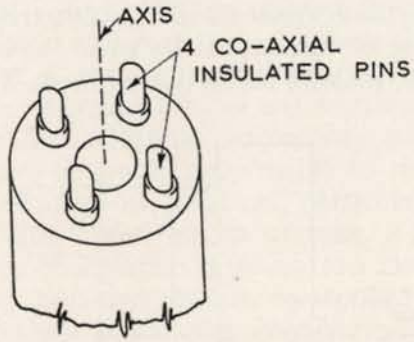


FIG. 5.1. *Four-pole electrostatic cylindrical lens for the correction of astigmatism.*

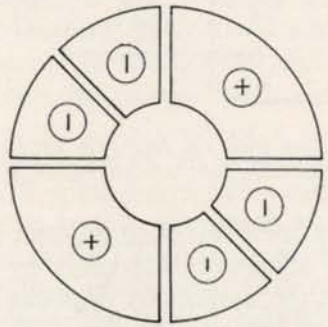


FIG. 5.2. *Schematic diagram of six-pole astigmatism corrector system allowing electrical rotation.*

special coupled potentiometers to give voltages proportional to the sine and cosine of the setting angle. This system therefore results in little overall simplification.

There is only very restricted space available in the region of the lens pole pieces where the object and the limiting aperture have to be positioned, so that the necessity of

locating the astigmatism corrector also in this region is an embarrassment. Haine and Mulvey (1954) have shown that the corrector can be positioned a considerable distance below the lens without affecting its performance. This is of great advantage in allowing more flexibility in design, but has two minor disadvantages: a small amount of image distortion results, and it is necessary to provide means for aligning the corrector with the objective lens axis, otherwise an image movement occurs when the corrector is varied in strength, making correction difficult. The image distortion

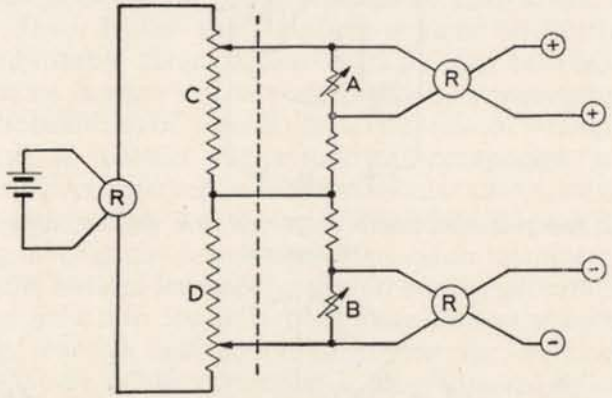


FIG. 5.3. *Circuit for use with the four-pole astigmatism corrector with provision for electrical centering.*

is quite negligible for normal amounts of correction, and alignment can be achieved electrically by applying a small potential between the diagonally opposite electrodes of the corrector as well as between the diagonal pairs. The circuit used to achieve this is shown in Figure 5.3. The method of centering is simple: having set the orientation roughly, the amplitude is increased to its maximum (maximum applied voltage), by the ganged control potentiometers C and D, and the supply voltage is switched off and then on again. The image on the microscope screen will be observed to move, but can be brought back to the position it occupied with the voltage off by adjustment of the potentiometers

A and B. Once this has been done, the position of the image remains independent of the amplitude of correction applied, unless a large change in orientation is made when a further correction for shift may have to be made.

Methods and Tests for Correction

The availability of a device for the correction of astigmatism is by no means sufficient in itself, since the state of correction is not at all easy to recognise in the image. It is therefore desirable to have a critical test which allows a ready recognition that adequate correction has been achieved. In principle, it should be possible to recognise the presence of astigmatism in an image, particularly if it is observed as the focus varies, but, in practice, a method based on such direct observation is much too insensitive to achieve anything but the crudest correction, except under special conditions where the observation of the Fresnel diffraction fringe is used, as will be described below.

Four test methods have been proposed for the detection of astigmatism, though only one of these, the Fresnel diffraction fringe method, has proved adequately sensitive and really practicable in the normal electron microscope. The methods are described briefly in the following paragraphs.

(*a*) *Beam rotation method.* In this method, due to Le Poole (1949), the illuminating beam is caused to rotate in a hollow cone with its apex at the object by a system of four pairs of deflector plates, each pair excited with an a.c. voltage in quadrature phase with that on the other. The object is thus illuminated with a hollow conical beam of electrons so that, if slightly out of focus, each point will be imaged as a circle. If astigmatism is present the circle reduces to a straight line when focused in either of the astigmatic planes and to an ellipse further from focus. Correction is achieved when the straight line is reduced to a point. The method cannot be satisfactorily used in the presence of an objective aperture, and although very quick and simple to use for a rough test, it is not sensitive enough

for correction down to the ultimate limit set by spherical aberration and diffraction.

(*b*) *Projection method.* This method, due to Castaing (1950), requires the placing of the objective lens in the position of the final projector lens in the microscope. An object, consisting of a fine straight wire stretched across the axis, is then placed after the lens. The image of this wire is distorted by the spherical aberration and astigmatism and from the distortion the astigmatism can be determined. The method, while sensitive, is not suitable for correcting an objective lens under normal operating conditions.

(*c*) *Caustic method.* This method, due to Grivet, Bertein and Regenstreif (1949), makes use of the properties of the caustic produced by the aberrations, to determine astigmatism. The lens is illuminated with a near parallel electron beam and the caustic produced near the back focal plane is observed after enlargement in the subsequent lenses. The method is potentially very sensitive but no way has yet been described of applying it to the microscope while the objective is operating under normal conditions.

(*d*) *Fresnel diffraction fringe method.* This method, originally suggested by Hillier and Ramberg (1947), fulfils the requirements that it can be carried out under operating conditions exactly coinciding with those for normal image formation and has adequate sensitivity. A full analysis of this method and its application has been described by Haine and Mulvey (1954). The method is based on the first (or, more correctly, zeroth) order diffraction fringe which appears in an out-of-focus image of the edge of an object. In Figure 5.4, A is the edge of an object illuminated with a near parallel electron beam. In a plane B distant z from A, and conjugate with the microscope screen, a diffraction pattern appears due to the interference between the direct wave from the source and the edge scattered wave from A. The optical path difference is given by:

$$\Delta = A'B - AB = A'B(1 - 1/\cos\theta)$$

where θ is the angle between AB and the direct ray direction A′B.

Intensity maxima will appear when $\Delta + q\lambda = n\lambda$, where $q\lambda$ is the path difference due to any phase change between the illuminating wave and the scattered wave produced by the object. Thus:

$$z(1 - 1/\cos\theta) + q\lambda = n\lambda$$

or

$$\tfrac{1}{2}/\theta^2 + q\lambda = n\lambda \quad (\theta \text{ small})$$

or y, the radial position of the fringes is given by:

$$y = z\theta = \sqrt{2z\lambda(n-q)}.$$

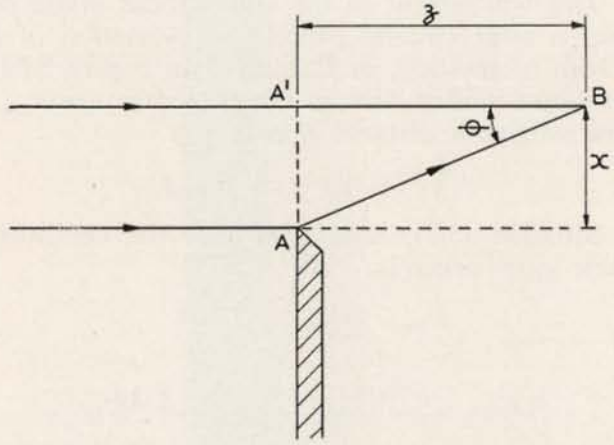


FIG. 5.4. *Diagram illustrating the derivation of the width of the zeroth order Fresnel diffraction fringe.*

The value of q would be expected to vary with the optical thickness of the object at and near the edge. According to the optical theory (viz.: Hillier and Ramberg, 1947), the distance, measured in the equivalent image plane, between the zero and maximum of the first fringe in the diffraction pattern should vary between $0{\cdot}45\sqrt{\lambda z}$ and $0{\cdot}7\sqrt{\lambda z}$ depending on the optical thickness near the object edge. In their experiments, Haine and Mulvey have shown that this spacing in fact lies between $0{\cdot}9\sqrt{\lambda z}$ and $1{\cdot}4\sqrt{\lambda z}$,

having a most likely value of about $\sqrt{\lambda z}$. There is no obvious reason for this discrepancy, but it may be due to the variation of scatter amplitude with angle. A further unexplained phenomenon is the marked asymmetry between the diffraction pattern obtained in under- and over-focused images. However, for the purposes of the test, the discrepancies are not serious. The mean value of $y_0 = \sqrt{\lambda z}$ can be used without serious error.

The presence of astigmatism has the effect that the focal plane for edges orientated in one astigmatic plane are longitudinally displaced by a distance z_a (the astigmatic distance) from the focal plane, for perpendicularly disposed edges. This will appear in the out-of-focus image of, for example, a large circular object as a variation of fringe width with orientation, as illustrated in Figure 5.6 (b). If the maximum and minimum fringe widths are y_{max} and y_{min}, the astigmatic distance is given by:

$$z_a = (1/\lambda)(y^2_{max} - y^2_{min}).$$

From equation (5.1), it is seen that the condition for negligible astigmatism is:

$$y^2_{max} - y^2_{min} < 1{\cdot}4d^2$$

or

$$(y_{max} + y_{min})(y_{max} - y_{min}) < 1{\cdot}4d^2. \qquad (5.2)$$

One very important aspect of the test can be seen from this equation. If $(y_{max} + y_{min})/2$ or the mean fringe width is large compared with d, the resolving power, then the second term becomes small compared with the resolving power and is difficult to measure. Also, by the same argument, the fringe asymmetry becomes very much less as the fringe width increases or as z, the distance off-focus, increases. Thus, the appearance of a symmetrical fringe is not, *per se*, an indication of adequate correction. The magnitude of the fringe width must also be taken into account. The net result is that the test must be performed very close to focus where the fringe width is comparable with the resolving power. The most obvious asymmetry is

obtained when the mean fringe width is about equal to the resolving power.

The Practical Realisation of Astigmatism Correction

In Figure 5.5 is shown a 'through-focus' series of an edge of a thin carbon film magnified 230,000×. It will be

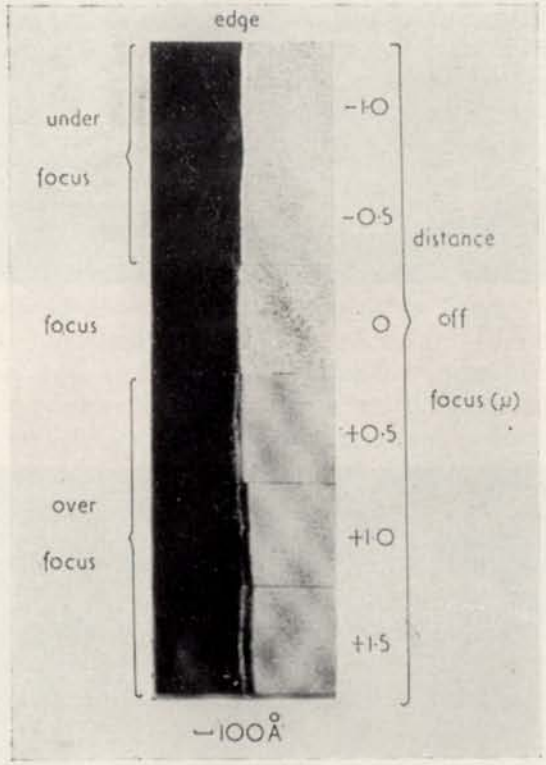


FIG. 5.5. *A through focal series in the image of an edge showing the variation of diffraction fringe with focus.*

noted that the over-focused image (lens too strong) is much more contrasty than the under-focused one. For this reason, the over-focused image is used in the test. Measurements of y are taken from the centre of the black line to the

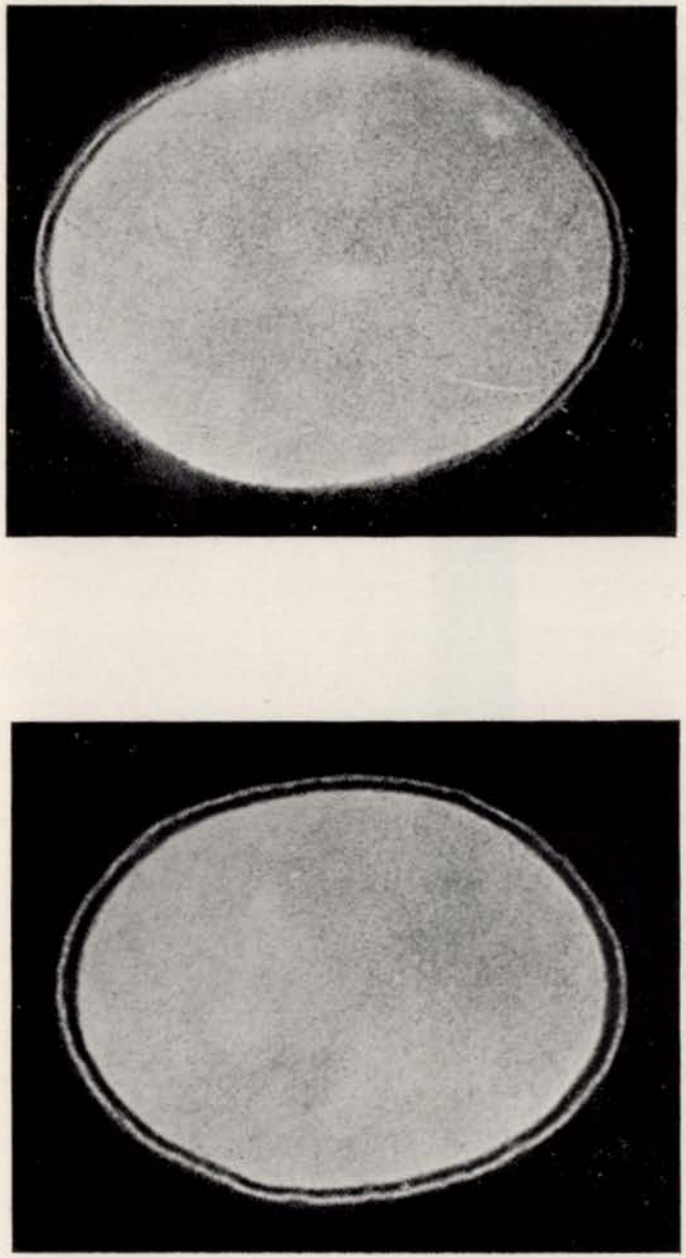


FIG. 5.6. *Image of a circular hole in a carbon film; (a) 1·4 microns of focus, and (b) with one direction of focus, illustrating the rapid increase in asymmetry as focus is approached.*

centre of the white line. Figure 5.6 shows a typical fringe obtained (a) 1·4 microns off-focus, and (b) with one direction on focus. The rapid increase of asymmetry as focus is approached is well illustrated here. The amount of astigmatism present is about that which may be expected from a well-made lens with no correction applied. The object is a hole in a thin carbon film prepared by the method described by Bradley (1954). This type of object is the nearest to the ideal one for the purposes of correction, and it is possible to support objects on the films so that correction can be applied without changing the specimen.

There are various factors which are of the greatest importance in assuring success in the application of the correction method, especially for high resolution working. Haine and Mulvey (*ibid.*) showed that the imagery of the Fresnel fringe is limited by all the factors, other than geometric aberration, which limit the normal resolving power of the instrument. On the other hand, it has been pointed out already that it is necessary to observe the fringe so close to focus that its size is comparable with the resolving power. It therefore follows that the test can only be applied to reduce the astigmatism sufficient for a resolving power d, if all other instrumental factors are themselves adequate for this resolving power. What is possibly more important is that, if the test can be successfully applied and fringes of width y obtained, then it is also shown that all other instrumental factors are adequate for a resolving power at least as good as y. Thus, integral with the test for astigmatism is the test for all these other factors. These include mechanical and electrical instabilities, etc., and are discussed later in this chapter. Since these factors, together with astigmatism, limit the resolving power of all but a very few instruments so far built to a value higher than that set by spherical aberration and diffraction, this test for resolving power is of the utmost significance and value, and the conditions for its successful application need careful consideration. Before discussing these, however, some of the other factors limiting resolving power will be discussed, since, as has been mentioned above, they are all

relevant to the test method, and the test method is relevant to them.

Instrumental Instability

It is clear that, if a high resolution image is to be successfully recorded, it must be stationary on the photographic plate at least within the limits of the resolving power. The following disturbances are the main causes of image movement:

1. Vibration of the specimen holder or mechanical stage with respect to the objective lens.
2. A slow unidirectional movement or drift of the specimen grid, holder or stage.
3. A flexing of the instrument column excited by vibrations.
4. Movement due to varying transverse magnetic fields.
5. Movement due to high voltage or lens current instabilities coupled with instrument misalignment.
6. Movement due to varying static electric charges on surfaces in the region of the electron beam.

Some of the detailed design requirements relevant to the instabilities above outlined are described in Chapter VIII. The effects of all of them are to produce image movement and spoil resolution. We are here interested in methods of detecting the source or sources of such effects.

If astigmatism and the chromatic defect arising from voltage and current instabilities are added to this list, it gives all the usual limitations which prevent the attainment of a resolving power in the region of 3–5 Å. To them must, of course, be added the effects of the specimen which will be discussed later.

Conditions for Optimum Viewing of the Fresnel Fringe and the Applications of the Test Method

It is desirable that the testing of the instrument should be carried out as far as possible by direct viewing of the

image on the fluorescent screen and without the necessity to resort to photography. Unfortunately, this is not entirely possible when the highest resolving powers are sought.

The resolution of the fluorescent screen is several times worse than that of the photographic plate. Hinderer (1942) has shown it to be about 70 microns. Agar (see Chapter VII) confirms this result when he shows that the Fresnel fringe contrast rapidly falls when displayed on the screen with a width less than 70 microns. It is an important advantage of the fringe test method that contrast variation along a line is much more visible than between a point and its surroundings. This results from the ability of the eye to integrate over a distance in one dimension.

To observe a fringe of width in object space d (Å) it is therefore necessary to magnify by $70 \times 10^4/d$. On the other hand, the image brightness falls as $1/M^2$, and the resolution of the eye and the observable contrast both deteriorate as brightness falls, and even with full dark adaptation it is not possible to obtain as bright an image as is necessary to see fringes as narrow as 5 Å. Some compromise is therefore necessary. In any case, an optical magnifier of at least ten times is necessary to see the 70-micron fringe on the fluorescent screen. Fortunately, as is described in Chapter VII, it is possible to magnify the image from the fluorescent screen up to about 20 × without loss of brightness, so that the fringe can be enlarged to 0·14 mm for the eye. This is just about adequate for the eye at reasonable brightness levels, such as can be obtained with a good electron gun system and with magnifications up to about 30,000 ×. This is suitable for a fringe width of about 20 Å.

It is possible to detect an asymmetry (δ) in the fringe of about 10%. Rearranging equation (5.2), the corresponding resolving power at which astigmatism will be negligible for a fringe of mean width y_m and asymmetry δ is given by:

$$d = 1{\cdot}2y_m\sqrt{\delta}. \qquad (5.3)$$

For $y_m = 20$ Å and $\delta = 0{\cdot}1$, d becomes 7·6 Å.

With practice, better correction can in fact be achieved

by the 'bridging principle'. As the astigmatism corrector is moved through the correct setting, the fringe asymmetry changes from one direction to a direction at right angles. The limiting settings where asymmetry is detectable in each direction can be noted, and then, by setting the control half-way between, the correction will be up to five times lower than the amount detectable. This same principle also allows an image to be focused to greater accuracy than can be seen on the screen. The correction requires two adjustments—one for amplitude and another for angle. To adjust the angle, in the first instance the direction of the fringe asymmetry with no correction applied is observed, and the amplitude of correction is then adjusted to a large value swamping the residual astigmatism to be corrected. Then the angular position of the corrector is adjusted to give an asymmetry as near perpendicular to that of the residual as can be achieved. This process is helped by switching the corrector 'off' and 'on' to observe the angle of asymmetry with and without the corrector. By following this procedure, the correction amplitude is reduced and should go through a setting where correction is achieved. In practice, the angular setting is usually not accurate enough to allow this. The small residual error in angular setting can be detected by watching the fringe asymmetry as the correction amplitude is reduced. If the angular setting is correct, the asymmetry will decrease to zero (symmetry) and then, on further reduction, will increase in a perpendicular direction. If, however, the angle setting is in error, the asymmetry will decrease and near correction will rotate from one direction to the perpendicular direction without reaching zero. The rotation of the asymmetry angle is easily noted and small adjustments of the corrector orientation can be made to reduce the asymmetry rotation effect to zero. It will be noted that this fringe asymmetry rotation effect is a dynamic one occurring as the correction amplitude goes through the correction value. Thus, the effect must be observed while 'rocking' the control around this value. It must further be noted that the optimum condition for observation depends upon the optimum setting of the focus

control, magnification, and image intensity (condenser focus). Only practice can make the operator proficient at this correction process, but it is important to know that with a well-adjusted instrument with no incidental defects, correction to below 5–10 Å can regularly be obtained by visual observation alone and without resort to photography. Figure 5.7 is a typical example.

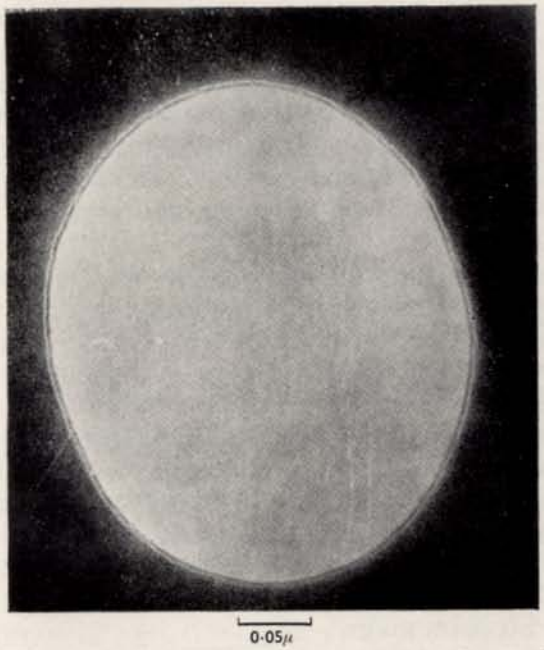


FIG. 5.7. *Typical edge diffraction fringe photographed after visual correction of astigmatism. The astigmatism deduced from remanent asymmetry is 0·1 micron giving a resolving power of 6 Å. (Magnification 185,000×.)*

The Overall Test Procedure

The regular attainment of high performance from an instrument can only be assured by the regular application of test procedures. Just how regular and how rigorous these procedures should be depends upon the basic quality of the instrument and just how exacting a performance is required. In the following, it will be assumed that a high-

quality instrument is available with the potentiality of giving a resolution better than 10 Å. The methods discussed can, in general, be applied to a lesser instrument, with modifications which will be readily apparent. The recommended test procedure is now outlined:

1. The instrument is aligned according to the procedure described in Chapter VIII.

2. A specimen consisting of a thin carbon film perforated with small holes is inserted. The instrument is operated at a magnification between 20,000–30,000 × in a darkened area (preferably completely dark or illuminated with a dim red light). Dark adaptation of the operator is necessary. A suitable circular hole is found measuring 1–3 mm diameter on the screen. This hole is observed through a 10–20 × optical viewing telescope of high aperture.

3. The illumination alignment is checked, preferably including a check on the gun brightness.

4. The freshly cleaned objective aperture is aligned (Chapter VIII).

5. The image is observed while varying focus. A dark fringe should appear inside the edge of the hole with slight over-focus, and with a defocused condenser lens. The setting of the condenser lens will need constant attention throughout the tests to find the best compromise between fringe sharpness and image intensity. The fringe becomes sharper the more the condenser is defocused, but the intensity becomes lower.

6. Having established the presence of the fringe, the focus is varied towards true focus; the fringe width decreases and the asymmetry increases.

As the defocus is further reduced, the fringe disappears. The disappearance may be due to the screen and eye limitation or may be due to the presence of instrumental defects. With experience, it is usually possible to judge whether the fringe disappears before reaching the limit set by screen or eye. At first it is desirable to resort to photography to check this. The focus is adjusted to a point just beyond the disappearing point. If before then the asymmetry is marked, it should be corrected by the astigmatism corrector. A photograph is taken to check whether the fringe is in fact still present. If so, the disappearance is due to screen and eye limitations. If not, an instrumental defect is present. In the latter case, this defect needs to be located (see later).

7. If the photograph shows a fringe, it should be between 10–20 Å in width, and, from its appearance, various deductions can be made. From the asymmetry, the astigmatism can be measured. This will be very small if the correction has been satisfactory. An asymmetric blurring of the fringe indicates an image movement due to one or other of the previously mentioned causes. In such a case, the minimum recordable fringe will vary with orientation. A wider fringe will tend to be 'blurred' asymmetrically. Such asymmetric blurring makes the accurate correction of astigmatism by the method described very difficult, as it is not easy under the difficult viewing conditions to differentiate between the asymmetric appearance of the fringes due to a variation in true width and that due to asymmetric blurring. The limit set to the degree of astigmatism correction, possible in the presence of such blurring, is just about equal to the limit of resolving power caused by the image movement. Again, therefore, the possible astigmatism correction is adequate for the resolving power limited by other causes. The resolving power of the instrument is *at least* as good as the recorded fringe width.

8. If no defects are detectable in the recorded fringe, the procedure may be repeated at a higher magnification. Alternatively, it may be repeated with an attempt to obtain a photograph even closer to focus which will give a narrower fringe more sensitive to the various defects. It will be recognised that the change of focus from that at which the fringe is still just visible, to a value near enough to focus to give the desired fringe width, is an operation which can only be judged by experience. It is of assistance, too, if the control is turned through focus until the underfocused (white) fringe is visible. This gives some indication of the sensitivity of the control. A better arrangement is to have means of

measuring the current change in the objective lens when adjusting the focus. The required change in focus is given by:

$$\delta z = (y_1^2 - y_2^2)/\lambda$$

where y_1 is the fringe width before the adjustment, and y_2 the width required. The fractional change in focus current is related to δz by:

$$\delta I/I = \tfrac{1}{2}\delta z/f.$$

Hence:

$$\delta I/I = \tfrac{1}{2}(y_1^2 - y_2^2)/\lambda f. \qquad (5.4)$$

9. From such observations, the following information is obtained:

(a) The fringe width asymmetry gives the residual astigmatism still requiring correction.
(b) The minimum recordable fringe gives the resolving power as limited by instabilities. Asymmetry in sharpness indicates an asymmetric limiting effect.

It should be noted that an experienced operator skilled in the use of these test methods can usually judge without photography whether a defect is present. Thus, if the fringe remains visible down to the limit set by the screen and eye, and the astigmatism correction is definite, the resolving power will be found to be at least as good as $(70/M)$ microns, where M is the magnification on the screen.

10. If a resolving power defect is found, it remains to find the cause of it. This can be a long and tedious job, and it may ultimately be necessary to call in the manufacturer's experts to perform it. Much can, however, be done to eliminate the probability that the defect lies in any of the classes outlined on page 100. Of course, much depends on how well the operator 'knows' his instrument, and on the confidence he can put into the design from various points of view. For example, if he can say 'I know a fault of such and such a nature is extremely unlikely in this design', he is that much better off. He is also fortunate if only one

limitation exists. A combination of two or more defects can be very much more difficult to trace. In the following, comments will be made relevant to detecting the presence of faults of the Type 1–6 previously referred to on page 100.

(i) *Vibration of the specimen holder or mechanical stage.* In this and several of the following defects, one is interested in detecting an effect which, as far as the instrument performance is concerned, is right on the limit of detectability. One general principle can be applied if a linear or near linear relation between cause and effect is assumed. For example, the vibration will presumably be excited by the general 'ambient' vibration of the instrument induced from sundry sources of vibration. If now the instrument is forced to vibrate at a known large amplitude giving an easily measured effect on resolution, and if the 'ambient' vibration of the instrument is measured, it may be assumed that the loss of resolving power due to the ambient vibration is that observed for the forced vibration multiplied by the ratio of ambient/forced vibration.

Unfortunately, complicated and not readily available apparatus is necessary for such tests. However, the sensitivity of the instrument to outside vibration can be assessed simply by tapping it with the finger tips. Such a test can at least tell whether the instrument has developed an increased sensitivity to vibration since a previous occasion when the resolving power was known to be good. Such a method depends on operator experience and/or training, and suggests an operator would be wise to train himself as to the sensitivity to finger tapping. The test is carried out while observing a high contrast object at high magnification through the viewing telescope, preferably a Fresnel fringe.

It should be noted that the part of the instrument which is causing the trouble may not be the specimen holder, though this is the most likely source. A loose magnetic shield has been known to give similar trouble.

(ii) *Drift.* Image drift can usually be detected simply by watching the image over a period of a few minutes against a reference mark on the screen. If it drifts one centimetre in three minutes, it can be expected to drift at about 50

microns per second, and the effect on resolving power over a two-second exposure can be calculated.

The cause of drift can be purely mechanical or thermal. Thermal drift may be due to the beam energy or to warming up of the objective lens. A drift test, as above described, with the beam 'on' all the time and with it 'on' only at the beginning and end shows the effect of the heating due to the beam. Usually, in a well-designed instrument, the main cause of drift will be due to beam heating. Very often a marked image movement will occur if the beam intensity is varied, due to thermal distortion of the holder or grid. This can usually be eliminated by ensuring good thermal contact between grid and holder, and holder and stage.

Thermal drift is greatly reduced if the total beam energy is reduced; this is achieved in high performance modern instruments by using very small areas of illumination obtained by double condenser lens systems or by screening apertures (Chapter VIII).

(iii) *Column flexing*. The general rigidity of the column can be tested by applying external forces by hand or by gently rocking the instrument, while observing the image. Flexing, if present, is usually due to an inadequate mechanical joint between sections. This can be located by 'squeezing' such joints with the hands and looking for any undue image movement.

(iv) *Transverse a.c. magnetic fields*. This is a very likely cause of image movement. The effect is usually unidirectional or asymmetric. A given instrument will lose resolving power if the ambient a.c. magnetic field exceeds some limiting value, usually between 0·1–10 gauss depending on how good is the magnetic shielding. The ambient field can vary due to the installation of new transformers or motors, etc., in the vicinity of the instrument or due to faults developing on wiring giving non-coincident 'go' and 'return' paths to a.c. current. It is useful to obtain information on the maximum allowable field from the manufacturer or by measurement (using the linear interpolation method already outlined for vibration), and to have equipment available for measuring the ambient field. A spool of fine

wire of about 200,000 turns about 6 cm mean diameter and a normal oscillograph with a sensitivity of a few millivolts provides adequate test facility.

(v) *High-voltage and lens current instability*. The effect of high-voltage and lens current instability worsening resolution by defocusing has already been discussed in Chapter III. Another cause of loss of resolution due to high-voltage or lens current instability occurs if the electron optical instrument is not perfectly aligned. Any variation in the high-voltage or lens current will then result in a lateral movement of the image (see Chapter VIII). It is not always possible in a defective instrument to produce an alignment which does not result in some residual image movement when either the lens current or high-voltage change; often alignment can be perfect for only one or other of these. Let us suppose that perfect alignment has been obtained for lens current, that is to say, the alignment has been arranged so that no lateral movement occurs even for large variations in the objective lens current. It may now be found that a variation in the high voltage does produce lateral movement. To ascertain whether this movement is significant it is useful to be able to make a controlled variation in the high voltage, say, by a fraction $\delta V'/V$. If this produces an image movement $\delta Y'$, then the image movement produced by a spontaneous variation in high voltage $\delta V/V$, the alignment error, is such as to produce an image movement, and hence loss of resolution of:

$$\delta Y = \frac{\delta V}{\delta V'} \delta Y'. \tag{5.5}$$

It is possible to use this principle to test the high-voltage or lens current stability. If a deliberate small gun tilt is introduced and a small known change in objective current ($\delta I''$) is made it would produce an image shift (say $\delta y''$). If there is a residual fluctuation of the lens current or high voltage, under the tilted gun condition, the Fresnel fringe will be blurred in one direction due to the resulting movement. If the estimated extent of this blurring is $\delta y'''$ an

indicated fluctuation in the high voltage or lens current is given by $\frac{1}{2}\delta V/V$ or $\delta I/I = \delta y'''/\delta y''$. Substituting this value of δV back in equation (5.5) gives the value for the limiting effect on resolving power. It will be noted that, in this test, by introducing a deliberate misalignment the effect of the high-voltage, or lens current instabilities, has been greatly exaggerated and has thus been made easily detectable.

(vi) *Static charges*. The effect of static charges building up on surfaces of insulators or on insulating surface films can cause slow or rapid image movement. Insulator surfaces should be well screened from the electron beam, both so that they cannot collect from it even scattered electrons, and so that if they become charged the resulting potential fields cannot influence the electron beam. The provision of adequate insulators and the associated screening arrangements is largely a matter of good design and can be presumed to have been adequately dealt with in a reputable instrument. Nevertheless, it does happen that contamination with insulating materials leads to trouble of this type. Very often the effect is to produce an instability of the image and a movement which is a function of beam current, and often magnification. The former can be confused with movement produced by thermal effects. The approximate position of the contamination can often be deduced by noting the effect of changing magnification, condenser lens strength, and of putting in or removing the objective or field apertures. The cure is to thoroughly clean the surfaces affected, or sometimes it is more convenient to cover the offending surface with a conducting layer of soot from a smoky flame.

Effects of Specimen on Resolving Power

However good its instrumental performance may be, the electron microscope image will not show high resolution unless the object is adequate. It is pertinent, therefore, to make some remarks in this chapter on the limitations imposed by the object. It was explained in Chapter IV how inelastically scattered electrons arrive at the object displaced

from the Gaussian image point, and thus produce a halo or fog around a single scattering point. The sum total effect of all the inelastic scatter in the object is to 'flatten' the contrast, just as that of a picture is flattened by viewing it through a ground glass plate displaced a short distance from it. This most important limitation, which can be imposed by the object, results from having too much matter in it. For this reason support films or replicas should be as thin as possible, as should the supported object.

One way of judging the adequacy of an object is to view the distribution of intensity in the back focal plane of the objective lens (see Chapter VIII). For fairly parallel illumination (high coherence, condenser defocused), this will show a very sharp central maximum, small compared with the object aperture. If scattering is excessive this sharp maximum will be blurred out to fill the whole aperture.

The absence of excessive inelastic scatter is not, of course, sufficient in itself to produce a good high resolution image. The object must also contain contrasty and fine detail. The provision of adequate test objects for the instrument has not been easy. It is for this reason that so much has been said in this monograph on the use of the Fresnel fringe for test purposes. Dispersion of fine aggregates of heavy metal on very thin carbon films make a reasonable test object, though it is never quite certain whether the resolution obtained in the image is limited by the object or the instrument, this is, in fact, generally true of any known object. All one can say of the result is that the instrument resolving power must be at least as good as the resolution indicated by the contrast variation in the image, but that it may still be limited by the object. The Fresnel method is much more certain in this respect. Figure 5.8 shows a micrograph of gold palladium alloy which has been evaporated on to a thin carbon film and then recrystallised by heating in the electron beam. This sort of picture may be obtained with good regularity on an instrument which shows a resolving power better than 10 Å by the Fresnel method.

An effect which has serious consequences on image resolution is contamination of the specimen. This effect,

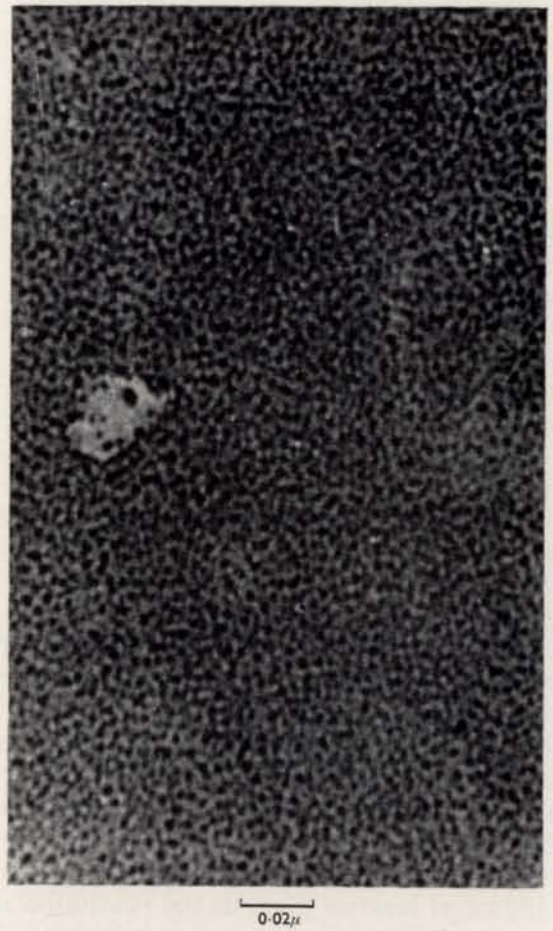


FIG. 5.8. *Fine particles of gold-palladium alloy.*
(*Magnification 500,000* ×.)

which has been studied by Ellis (1951) and Ennos (1953), results when the electron beam falls on the hydrocarbon molecules which condense on the specimen in the microscope. The sources of these molecules include the oil diffusion pump, rubber gaskets and also any oil or grease which may be present on metallic surfaces. Ennos (1954)

showed that only by fairly prolonged baking could the hydrocarbon vapours be reduced to an adequately low pressure. The molecules on the specimen surface appear to be in equilibrium with the vapour, at room temperature, so that some fractional monolayer is present. The effect of the electron beam is to dissociate the molecules on the surface to leave carbon which remains fixed to the surface. Further molecules then condensing are likewise affected, resulting in a continuous growth of the contaminating deposit. At normal beam intensities and with the object unheated the growth rate is about 1 Å per second. If the specimen is heated by the electron beam or otherwise, the equilibrium concentration of condensed molecules is lower, so that contamination is reduced. Thus at 200°C the rate of contamination is usually negligible. One solution of the difficulty is, therefore, to use a heated specimen stage.

For an illuminating beam diameter of 20 microns, as normally obtained with a single condenser lens, the object will normally operate at a temperature elevation sufficient to reduce contamination appreciably. However, this very heating can lead to temperature instability of the specimen holder and stage and hence image movement. For this reason two stage condenser lens systems have been introduced to give smaller illuminating spot diameters and less heating (Chapter VIII). The use of these can lead to the contamination effect becoming a serious limitation. Leisegang (1954) has described a method of reducing contamination in which the specimen is surrounded by a trap cooled to −70°C. The results showed the method to be effective but it is more complicated than raising the temperature to about 200°C and suffers from the same disadvantages. The principle of the method is to trap the molecules before they arrive at the specimen.

The contamination difficulty need not be a serious limitation provided the beam is not allowed to irradiate the object for more than a few seconds. This would be more easily achieved if the image intensifying system, described in Chapter VII, could be put into practice.

Conclusions

In this chapter an attempt has been made to outline the most important factors which may limit resolution to a value worse than that theoretically predicted. The remarks made are drawn mainly from the author's own experience and from many discussions with others in this field. It is inevitable that other limitations will arise from time to time and it should be stressed that the best performance can only be expected by an operator who is fully conversant with the design and working of his instrument, and who is familiar with the types of test procedure here described.

REFERENCES

BERTEIN, F. *Ann. Radio. Elec.*, **3** (1948), 379.

BRADLEY, D. *Proc. Internat. Conf. on Electron Microscopy*, London, paper 108 (1954), 478.

CASTAING, R. *Compt. Rendus*, **231** (1950), 835.

ELLIS, S. G. *Proc. Symposium on Electron Physics*, Washington (1951).

ENNOS, A. E. *Brit. J. Appl. Phys.*, **4** (1953), 101.

ENNOS, A. E. *Brit. J. Appl. Phys.*, **5** (1954), 27.

GRIVET, P., BERTEIN, F. and REGENSTREIF, E. *Proc. Conf. on Electron Microscopy*, Delft (1949).

HAINE, M. E. and MULVEY, T. *J. Sci. Instrum.*, **31** (1954), 326.

HILLIER, J. and RAMBERG, E. *J. Appl. Phys.*, **18** (1947), 48.

HINDERER, K. *Z. Phys.*, **119** (1942), 39.

LEISEGANG, S. *Optik*, **10** (1953), 5.

LEISEGANG, S. *Proc. Internat. Conf. on Electron Microscopy*, London (1954). Published by the Royal Microscopical Society.

LE POOLE, J. B. *Proc. Conf. on Electron Microscopy*, Delft (1949).

RANG, O. *Optik*, **5** (1948), 518.

CHAPTER VI

THE ELECTRON GUN CHARACTERISTICS

THE electron gun is the illumination source for the electron microscope. Its performance is a good deal more critical than the light source in the optical microscope, firstly because the much higher magnification used requires a much greater intensity of illumination, secondly because the chromatic aberration of the objective lens requires a very narrow wavelength spread and, hence, a high voltage stability, and thirdly because the spherical aberration of the objective lens limits the usable beam angle to a very small value requiring a very high beam brightness* or current density per unit solid angle.

We now discuss the requirements of an ideal gun to give adequate electron microscope performance. If the minimum usable fluorescent screen brightness is obtained with a current density ρ_s, the necessary current density at the object is $M^2\rho_s$, where M is the magnification from object to screen. The illuminating aperture angle α_c is normally required to be less than the objective aperture angle α_0. Let $\alpha_0 = \kappa\alpha_c$, then the required current density per unit solid angle or brightness (β) is given by:

$$\beta = M^2\rho_s\kappa^2/\pi\alpha_0^2. \tag{6.1}$$

The required magnification is given by the ratio of the size of an elemental image point δ, corresponding to the resolution area, to the resolving power d

$$M = \delta/d.$$

* The use of the term 'brightness' for current density/unit solid angle is often deprecated as not being accepted as a photometric unit. Nevertheless, it is widely accepted in electron optics and it is used here with apologies to those who dislike it.

Putting $d = 0{\cdot}6\lambda/\alpha_0$ and substituting in (6.1) gives:

$$\beta = 0{\cdot}9\delta^2\kappa^2\rho_s/\lambda^2. \tag{6.2}$$

Note that this is independent of α and therefore the resolving power. Hence, no better gun performance is required for an instrument of resolving power 2 Å than for one of 100 Å.

Now the current density per unit solid angle, or brightness, obtainable from an electron gun has a theoretical invariant value, limited by thermodynamic considerations. The theoretical value is given by:

$$\beta = \rho_c\phi_0/\pi kT \tag{6.3}$$

where ρ_c is the cathode current density, ϕ_0 the gun accelerating voltage, k is Boltzmann's constant ($8{\cdot}6 \times 10^{-5}$eV/°K) and T the cathode temperature in degrees absolute.

This expression has been widely misused in electron optics and its proof in the general case appears rarely in text-books. For these reasons some further discussion on it, and a fairly general proof, are given in an appendix to this chapter.

Equating this value to that required by equation 6·2 and putting $\lambda = \sqrt{\frac{150}{\phi_0}} \times 10^{-8}$ cm we get for the required cathode current density:

$$\rho_c = 1{\cdot}6 \times 10^{10}\delta^2\kappa^2\rho_s T. \tag{6.4}$$

It will be noted that this is independent of both resolving power and voltage.

Filament Life

The relationship between cathode current density and 'life' for tungsten filaments has been discussed by Bloomer (1957). The 'life' is determined by erosion from gas attack and by direct thermal evaporation. Gas erosion varies only slowly with temperature between 2500 and 3000°K. Figure 6.1 reproduces Bloomer's curves and shows the lives of 0·004 in., 0·005 in. and 0·006 in. diameter filaments as a function of pressure of air or water vapour which are

about equally destructive. The curves are for an operating temperature of 2900°K. For other temperatures, the lives indicated must be multiplied by the factor shown in the

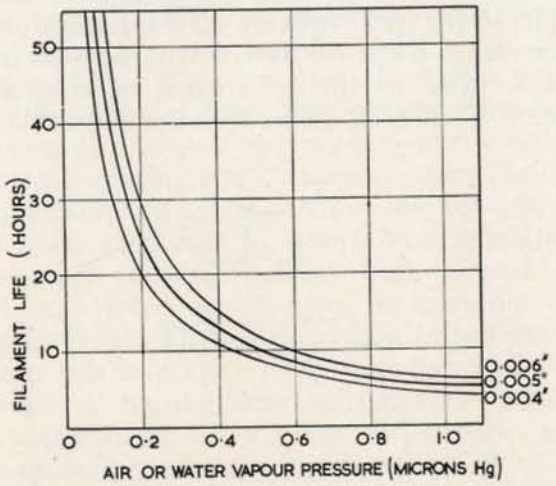


FIG. 6.1. *The variation of cathode life with air or water vapour pressure as limited by gas erosion for three different cathode wire diameters. Temperature 2900°K.*

table below, which also gives the corresponding current densities. The relatively slow dependence on temperature will be noted.

TABLE 6.1

Multiplying factor for oxygen or water vapour erosion as a function of temperature and emission density from a tungsten cathode

Temp °K	2500	2600	2700	2800	2900	3000
Factor	1·36	1·25	1·15	1·04	1·00	0·92
Current density (amp/cm²)	0·3	0·7	1·6	3·5	7·3	14·2

The values plotted are based on the assumption that one-third of the filament erodes away before life is ended.

The fact that the erosion rate is relatively independent of temperature results in the diameter reducing uniformly along the hot part of the filament. Thermal evaporation, on the other hand, is strongly dependent on temperature and results in the wire 'necking' at the hottest point. This difference makes it easy to differentiate between a filament which has ended its life by erosion as compared with evaporation. Figure 6.2 shows such a comparison.

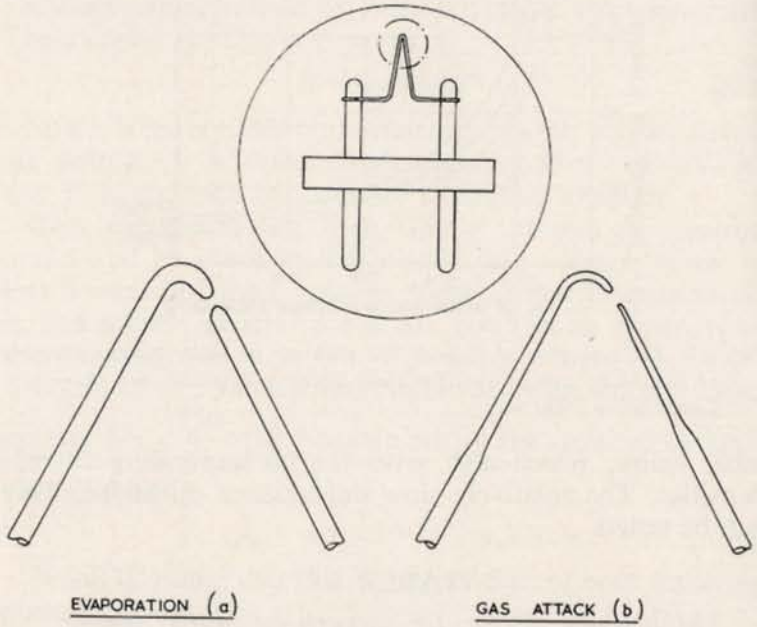


FIG. 6.2. *Appearance of filament wire after failure by evaporation* (a) *and gas attack* (b).

The evaporation process tends to be unstable, but the instability depends very closely on the electrical conditions of the heating current supply. In the two extremes, the circuit may apply a constant current to the filament or a constant voltage. In the former case, as soon as the hottest part of the filament begins to thin, its resistance will rise and hence the power fed to this part will increase, it will get hotter and the evaporation rate will increase and so on.

In the case of a constant voltage supply, the temperature will drop as thinning occurs. In any practical case on the electron microscope, the supply must inevitably approximate closer to the constant current case.

Bloomer computed the lives of filaments due to evaporation from vapour pressure data due to Reimann. The lives were computed for a 6% thinning at the hottest point. The result is little different from that obtained for complete failure, owing to the instability already mentioned, which results in the evaporation rate increasing very rapidly once the thinning exceeds a few per cent of the diameter.

The diagram of Figure 6.3 shows how brightness and life varies with heating current and temperature for 0·0048 in. and 0·005 in. diameter hairpin filaments of 0·55 in. wire length. The relationship with heating current is useful but only refers to a fairly wide hairpin of the wire length given. A sharply bent hairpin will operate at a different temperature, for a given current, owing to the shielding action between the legs near the tip. This also causes the filament to run at a lower temperature at the tip than at a point about a millimetre down each leg. For this reason, filaments usually burn out at the side instead of at the tip. This can be avoided, with some small advantage, by grinding a small flat on the tip.

Bloomer's results show that the life (t_l hours) is very nearly inversely proportional to the available current density according to the relation:

$$t_l = 32/\rho_c. \tag{6.5}$$

The value is only slightly dependent on wire diameter.

For a high performance instrument, a life of about 30 hours is quite acceptable, especially since a well-designed gun will allow a filament change to be made in 5 to 10 minutes. Thus, a current density of about 1 amp/cm^2 can be made available. In special circumstances, this might even be increased to 5 or 10 amp/cm^2.

It will be seen from Figure 6.1 that the life will not be limited to less than 30 hours by chemical erosion provided the vacuum is better than 10^{-4} mm Hg. Such a pressure is

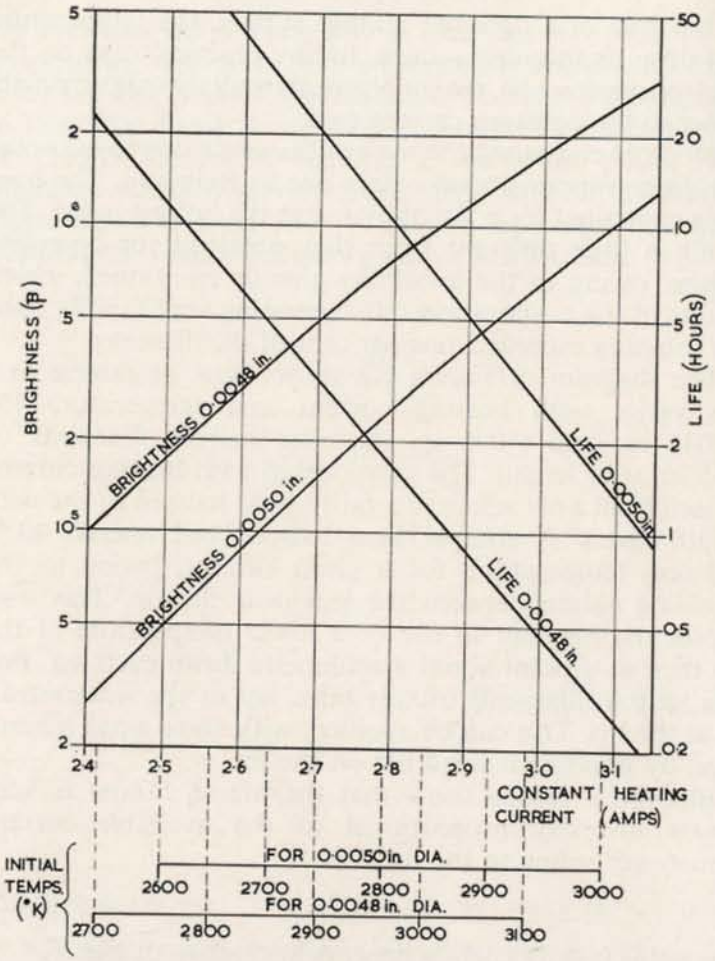


FIG. 6.3. *The variation of theoretical electron gun brightness and cathode life (as limited by evaporation) with heating current and temperature for initial wire diameters of 0·0048 in. and 0·0050 in. (14 mm long filaments).*

not difficult to maintain, but erosion trouble may be experienced in microscopes not fitted with specimen and camera airlocks. In this case, the repeated 'letting down' to atmospheric pressure results in a higher operating pressure mainly due to the adsorption of water vapour. It

is important, in this respect, to remember that the rate of adsorption is quite slow and the shorter the time at atmospheric pressure the longer will be the filament life. This effect is significant for times of up to 10 minutes at least.

Further reference will be made to the requirements of the electron gun and equation (6.3) in Chapter VII.

Practical Electron Gun Performance

Apart from the necessity for giving a beam brightness approaching the theoretical value, it is also necessary that the electron gun shall provide a beam of suitable total divergence angle, and one focusing into a suitable spot size. It is convenient for the gun to provide a beam of larger divergence angle than required and then to limit the beam, by a suitable aperture, to the required value. The advantages are that the gun normally gives a beam with a radial distribution of current density falling off from the centre in a roughly Gaussian manner: by selecting only the centre part of this, a higher total current is obtained in a beam of given divergence angle. Also, the alignment of the gun is less critical when an aperture is used with a large total beam angle, since the final beam position is determined by the aperture and not by the initial beam. This arrangement further makes the choice of the gun electrode geometry less critical, as will appear later in this section. Against these advantages must be set the disadvantage that a greater total current is required. However, the required beam current is never more than a few hundred microamperes and its provision should not present any special difficulty except, possibly, where it is desired to use an electrostatic generator for the high voltage supply.

The choice of a tungsten cathode seems more or less inevitable, at least for the present. Other cathodes of lower work function and operating temperature, such as the oxide-coated, impregnated or thoriated types, promise as high an emission but are not suitable for use in the rather poor vacuum conditions and at the high voltage found in the electron microscope. Attempts have been made to use a cold cathode gun in which the electrons are emitted by

secondary emission in a gas discharge (Induni, 1947). This, though claiming advantages in life, does not appear to give as high a brightness as the tungsten filament. Recently experiments have been described in which the field emission from a sharply-pointed tungsten wire have been used. Whereas it is true that field emission promises to give a cathode current density many times higher than the tungsten filament, the main advantages of the sharply-pointed field emission cathode appear, up to the present, to lie in the smallness of the source which is useful in certain cases where a very high coherence is required (Hibi, 1956).

The design of the gun in general use was evolved experimentally. It comprises a tungsten 'hairpin' filament of wire which may be 0·005 in. or 0·006 in. diameter. This cathode is mounted in a 'shield' which essentially comprises a disc, perpendicular to the axis, with a central aperture within which the tip of the cathode is positioned. The anode is a flat plate with an axial aperture through which the emitted electrons pass. Originally, the shield was connected to the cathode but it was found that considerable improvement resulted by applying a negative bias to this electrode. As the microscope improved it became more and more obvious that gun performance was of the greatest importance and investigations were made to find out more about its characteristics.

Early investigations were hampered by a lack of appreciation of the basically important parameters and were concerned more with obtaining the narrowest beam divergence angle rather than the highest brightness. A mistaken conclusion was reached that space charge effects were the main factors in determining the gun performance (Hillier and Ellis, 1949). The importance of the various characteristic parameters were later appreciated and investigations were pursued to evaluate the gun performance and its mode of operation. The most important requirement is the highest possible brightness consistent with reasonable life. Of incidental importance is the required total current which may affect the high-voltage generating unit; it may possibly be of importance also in relation to

the heat dissipation at the condenser aperture, and is relevant to the question of beam alignment before the condenser aperture, as already mentioned. The choice of filament diameter is determined, to a certain extent, by the required spot size and by the total heat dissipation from it. The required bias voltage is of some importance from a design point of view in that adequate insulation must be provided and a suitable source of potential found.

Haine and Einstein (1952) have described an extensive investigation on the electron microscope gun. The type of system studied is shown in Figure 6.4. It comprises a 'Vee'

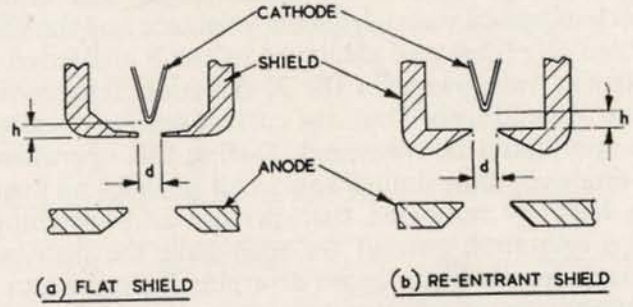


FIG. 6.4. *The flat and re-entrant cathode shield geometries studied by Haine and Einstein.*

tungsten filament bent from 0·005 in. diameter wire, a shield with an axial aperture which may be flat as in (a) or shaped as in (b), and a flat anode with a central aperture. The investigation was aimed at finding the optimum design and operating conditions which would give the maximum brightness, reasonable economy in total current, and at elucidating the mechanism of the beam formation in the gun. It will be noted that brightness and current are not necessarily directly related. For a fixed source size the brightness is proportional to the current which passes a small aperture placed in the beam. This current can be independent of the total current, which may vary due to a variation of the total beam divergence angle only. It should further be explained that by source size is meant the apparent source size from which the accelerated electrons

diverge. The size of this source, or virtual source as it may be more aptly described, may vary with the beam angle used in imaging it if geometric aberrations are present in the gun.

In the experimental investigation, a gun was used in which the height of the cathode behind the grid aperture could be varied in operation. This height h was defined as negative if protruding through the aperture. The cathode could also be centered whilst in operation. A magnetic lens was used to form an enlarged image of the virtual source and the beam could be swept, by magnetic deflectors, over a small aperture above a Faraday cage. The collected current was passed through a load resistance and the voltage generated amplified and displayed as the Y deflection on a cathode-ray tube swept in the X direction synchronously with the original scan. Thus, the current density distribution in the spot could be measured. During this operation the beam angle could be limited to a small angle by an aperture in the lens. By removing the aperture and switching off the lens excitation current the scan gave the distribution of current density in the beam diverging from the gun. The total beam current and bias voltage on the shield was also recorded.

With this apparatus, the variation of brightness (current density per unit solid angle*), total current, beam divergence angle and spot size could be measured as a function of applied bias voltage for different values of accelerating voltage, cathode temperature and for different geometries of the gun electrodes.

Considerable attention was paid to the question of temperature measurement of the cathode since emission and life vary rapidly with it. A cylindrical diode with guard rings and a cathode made from the filament wire used in the gun was built and run on a continuously pumped system. The emission was measured as a function of temperature measured on a disappearing filament optical pyrometer. The measured values agreed closely with

* The unit solid angle being defined by the aperture.

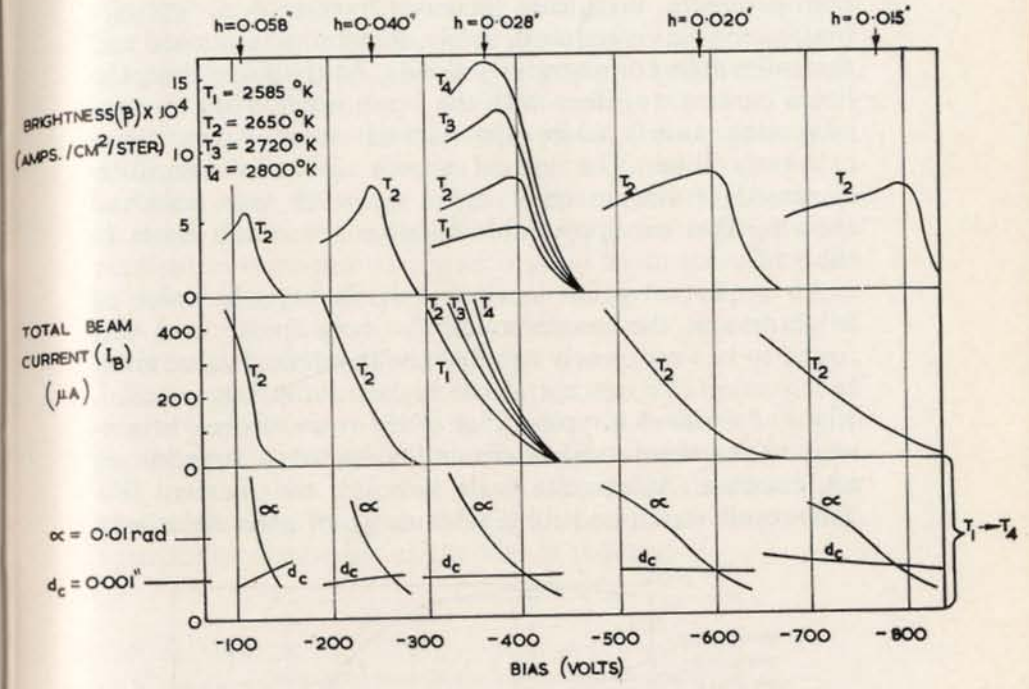


FIG. 6.5. *Brightness (β), total beam current (I_B), semi-beam angle (α), and virtual source diameter (d_c) plotted as a function of gun bias for five different values of filament height and for one height for four different cathode temperatures. Cathode shield aperture 0·07 in. diameter, accelerating voltage 50 kV.*

previously published values. This was held to confirm the pyrometer accuracy.

The filaments used were calibrated for temperature against heating current in a bell jar before use. The heating current was then used to set the temperature at the required value.

The results obtained in this work were enlightening. Figure 6.5 shows a typical set of characteristics for five values of filament height, but otherwise fixed geometry. Curves for brightness (β), total current, beam divergence angle and virtual source size, are shown for different temperatures as a function of bias. It will be noted that, at

each geometry, brightness increases from zero, at 'cut-off' bias, as the bias is reduced, until a maximum is reached and thereafter falls comparatively slowly. At the same time, the beam current, together with the beam angle, rises at ever-increasing rates, while the virtual source size varies extremely little. The virtual source size was also little increased if the imaging beam aperture was removed showing that no appreciable *imaging* aberration exists in the gun.

An important point in the results is that the value of brightness at the maximum of the brightness curve was found to be very closely equal to the theoretical value given by equation (6.3) except at the highest temperatures used. Figure 6.6 shows a typical plot of the ratio of peak brightness to its theoretical maximum value as a function of temperature. A separate scale indicates the filament life. This result was found for a wide range of geometries with

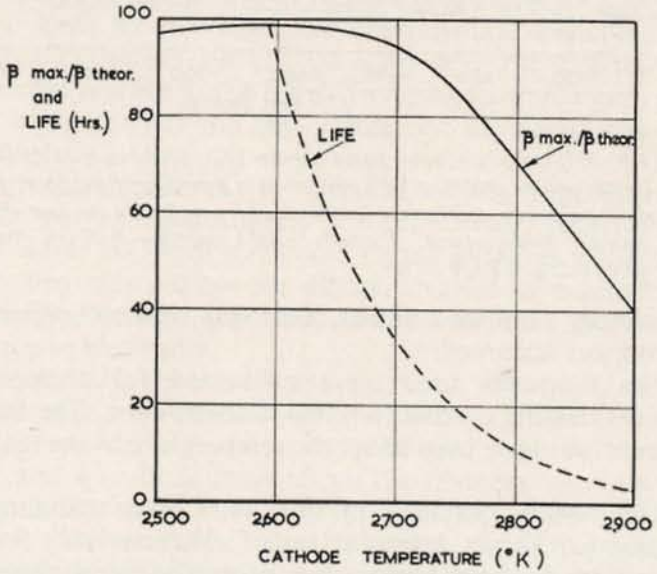


FIG. 6.6. *The ratio of the maximum brightness obtained to the theoretical brightness* ($\beta_{max}/\beta_{theor}$) *as a function of cathode temperature showing incidence of space charge effect. Plotted also is the cathode life.*

varying shield hole diameters, shield shape and height. The conclusion to be drawn is that geometry is relatively unimportant in obtaining the theoretical brightness, but, for any given geometry, the choice of optimum bias is important. A further conclusion, resulting from examination of Figure 6.6, is that there is no space charge limitation to emission over the range of cathode temperature likely to be used. The fact that the beam divergence angle is independent of temperature shows that no beam spreading due to space charge effects is present Thus, contrary to all previous belief, space charge effects were proved to be unimportant in this type of gun.

The experiments also showed that the axial position of the virtual source relative to the true cathode varied with the beam current and filament height. Figure 6.7 shows this variation for different heights for a flat shield and for one height for a re-entrant shield. It is seen that the position approaches the anode as the bias is reduced.

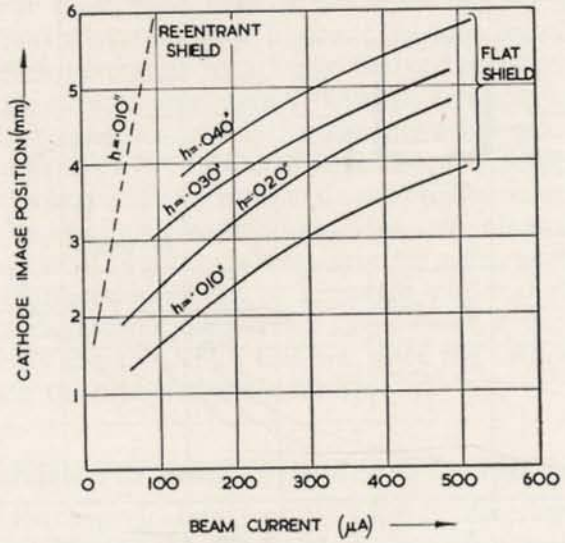


FIG. 6.7. *The variation of cathode image position with beam current for three values of cathode height with flat shield and one value of height for re-entrant shield (cathode aperture diameter 0·05 in., accelerating potential 50 kV).*

Examination of the distribution of current density across the beam, as the bias was reduced from cut-off, showed that, for the region up to that for maximum brightness, the distribution was near Gaussian. After this (less bias) the beam tended to become slightly hollow and, with still less bias, broke up into an irregular pattern. These effects have been explained previously as space charge effects. They were explained by Haine and Einstein as purely geometric effects, a fact confirmed by the observation that the beam structure was unaffected by temperature. An explanation will be given in a later paragraph.

Although the theoretical brightness can be obtained over a wide range of geometrical shape, the total beam angle and, therefore, the total current at which the optimum brightness is obtained, varies with geometry. In Figure 6.8

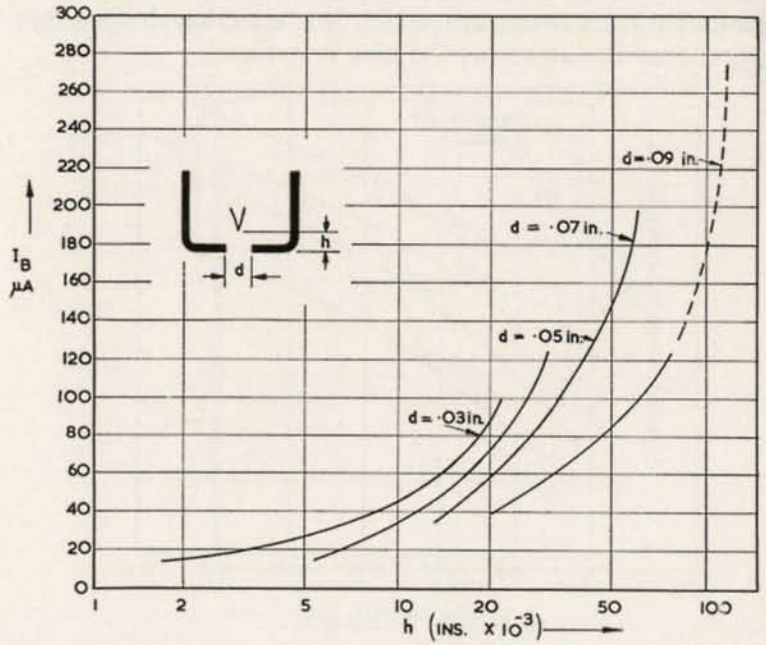


FIG. 6.8. *Variation of beam current at maximum brightness ($\beta = 80{,}000$ amps/cm²/ster. rad) with filament height for four values of cathode shield diameter.*

the value of the total current under optimum bias conditions (theoretical brightness) is plotted as a function of cathode height behind the shield for a number of shield aperture diameters. It is seen that smaller values of filament height, and larger shield aperture diameters give lower beam currents at optimum brightness (conditions giving higher bias voltage). Thus, for applications where only a small total beam angle is required, economy in total current is obtained by using reduced filament height and increased shield aperture diameter. The beam which would pass through the aperture defining the beam angle is quite unaffected by this economy. In the limit, the geometry can be chosen to give just the required beam angle, but it must be remembered that the current density distribution across the beam is of Gaussian shape. A higher current density in the focused image is, therefore, obtained if the beam angle is made several times the required value and the beam 'clipped' by an aperture to the required size. The whole aperture is then filled with a beam of the full current density, given by the peak of the Gaussian distribution. It is also of advantage to arrange the geometry to give a beam angle several times larger than the required value to reduce the required accuracy for the gun alignment. Thus, if the filament is a little off centre, the beam emerges as an angle to the axis and may not fall on the condenser aperture. A larger beam angle clearly allows a greater tolerance.

If a large total beam angle is required for some application then the filament height can be increased, and bias reduced. This procedure must, however, have a limit if the source size remains constant. This follows, since the cathode area and, hence, the total current is limited.

The Mechanism of Beam Formation in the Electron Gun

From the experimental results already described, and other considerations, it is possible to deduce a reasonably accurate picture of the mechanism of the beam formation in the electron gun. The experiments show that, within the usable range of cathode temperatures, space charge has

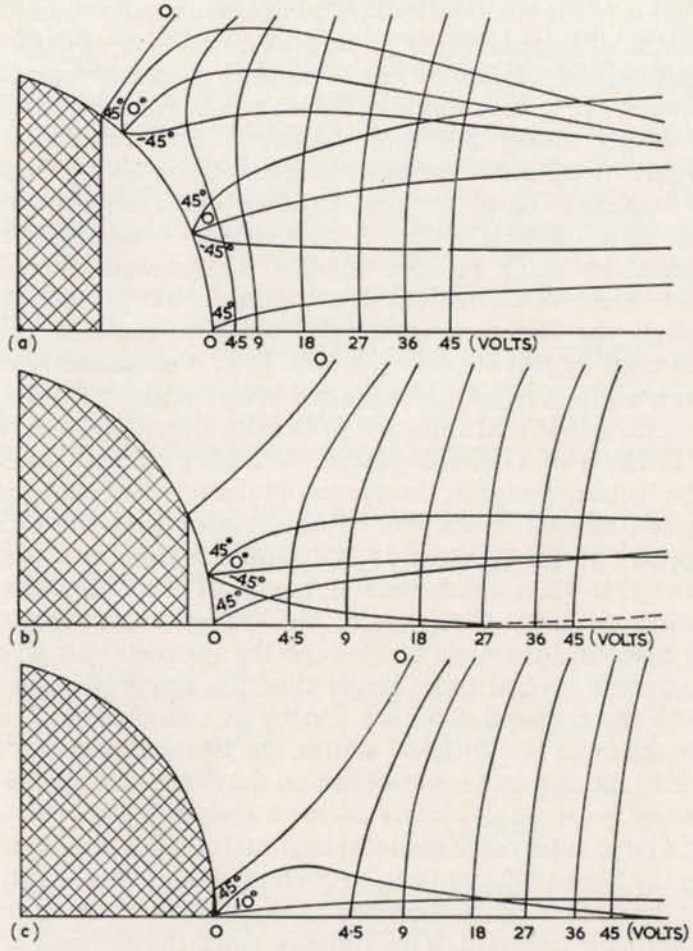


FIG. 6.9. *Potential field and electron projectory plots near cathode tip for three values of bias (initial velocity 0·3 volts).* (a) *At low bias* (I_B *500* μA). (b) *At optimum bias* (I_B *100* μA). (c) *Near cut-off* (I_B *10* μA).

negligible effect on the beam which, therefore, is determined purely by geometric considerations.

The bias voltage has little effect on the field distribution except close to the cathode. Figure 6.9 (a), (b) and (c) shows the distribution of field (by equipotential plots) in

the vicinity of the cathode tip. The zero equipotential which, at large radius, is between grid and anode, will intersect the cathode around a roughly circular line. Behind this line the gradient at the cathode is negative and no emission is possible. In front the field is positive and emission takes place. It will be noted that, near the axis, the field is fairly flat and the emitted electrons proceed approximately parallel to the axis, spreading due to the radial component of emission energies. Further out there is a strong radially inward component of field, which acts on the relatively slow electrons and guides them across the axis. This radial field increases rapidly with radius having a large component increasing with the cube of the radius; thus it acts like a fourth-order plate in close proximity to the cathode. The electron rays produce an aberration caustic in front of the cathode and somewhere, in their image, is a minimum cross-section. This is the virtual source. If only a narrow axial cone is selected by an aperture the image shows the draw marks on the tungsten wire. With wider angles these marks are lost since now the caustic cross-section is focused. For wide angles (low bias) the virtual source is well forward of the cathode towards the anode, as was shown in the experiments by the strength of the focusing lens required to produce a focus.

At high bias, near cut-off, the radial field distortion is quite close to the axis since the zero equipotential intersection is near the axis. Under these conditions the aberration causes the electrons with high radial velocity component to be focused quite differently to those with low radial velocity component. This explains the fall off of the measured brightness near cut-off bias. At some intermediate bias the effect of the radial field component is a minimum and the gun gives theoretical brightness.

As the bias is further reduced the zero equipotential begins to move round the sides of the filament tip, producing a very strong effect on the outer part of the beam which takes on the well-known hollow form. Still less bias exposes the back of the filament wire, leading to emission into a complex field geometry and the formation of a

complicated angular distribution in the beam and in the focused spot. The image splits up into a multiple pattern, always containing an odd number of spots—one for the tip plus an even number, as would be expected by the hairpin cathode symmetry. It should be stressed that this multiple spot pattern is still not appreciably changed by cathode temperature, even for a variation of total beam current by 100 times. This would seem to show conclusively that it is of purely geometric origin and not due to any space charge effect as has been suggested [Dolder and Klemperer, 1957].

Self Bias

Since the bias voltage to the electron gun must be supplied at 50–100 kilovolts above ground potential, it is usual to generate it from the potential drop in a resistance in series with the high voltage supply, Figure 6.10. The use

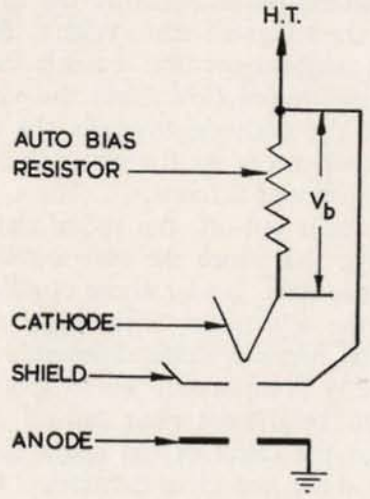


FIG. 6.10. *Circuit arrangement for automatic bias.*

of this self-biasing system leads to the imposition of further limitations on the electron gun design. The bias voltage (V_b) generated will now be proportional to the electron current

so that any increase in this current increases the bias voltage tending to counteract the current increase. The resistance thus introduces negative feed-back action which limits and stabilises the beam current. The stabilising effect is important. The emission current density varies exponentially with cathode temperature and hence is inherently not very stable. Any resulting current change can result in changes in the accelerating voltage unless the high voltage source impedance is very low. The stabilising action of the negative feed-back, due to the self-biasing arrangement, greatly reduces this tendency. For a given value of bias resistance the current will increase with cathode temperature up to some 'saturation' point, imposed by the negative feed-back action, a phenomenon well known to the electron microscopist, and at one time explained by space charge saturation.

For any given electrode geometry and bias resistance value, the saturation point will occur at some cathode temperature which is not necessarily the required operating temperature. Furthermore, the saturation point may not correspond to the point of maximum brightness. It is, therefore, important to consider whether a suitable combination of electrode geometry and bias resistance can be found where the saturation point occurs, both at the desired operating temperature and at the point of maximum brightness for this temperature. It is also desirable to choose a design giving this coincidence, together with a reasonable economy in total current. Consideration to this question has been given in a second paper by Haine, Einstein, and Borcherds (1958).

A restricted range of geometry only was considered. Of importance are two shapes of shield electrode as shown in Figure 6.4, one flat and one with a re-entrant cone. The filament height, temperature and bias resistor were varied and the total current and ratio of measured brightness to optimum brightness at each cathode temperature measured. The results are shown in Figure 6.11. The full lines show the currents as a function of temperature, or cathode heating current, for different values of cathode height

and bias resistance, while the dashed lines show the brightness ratio. The current saturation effect due to negative feed back is apparent and the saturation values are indicated by the dash-dot lines cutting across the current lines. It is desirable to operate in the region to the right of these saturation lines in the interest of beam current stability. It will be noted that the brightness ratio curves are independent of height for the flat shield but not for the re-entrant shield.

To find the design conditions for operation at a particular desired temperature (i.e., brightness and/or life) it is necessary to look for conditions, as shown by the set of graphs, where, for that temperature, the current saturation has been reached, the brightness ratio is near unity and the total current and beam angle have reasonable values. The beam angle is, of course, given by:

$$\alpha = \sqrt{I/\pi^2 r_s^2 \beta}.$$

Table 6.2 shows typical ranges of geometry and bias which might be used for various operating temperatures

TABLE 6.2

Optimum conditions for an electron gun under self bias

Shield	Fil. Temp. (°K)	Fil. Life (hr)	β_t theoretical (A/cm² ster.)	Bias resist. (MΩ)	Fil. ht. ($h \times 10^{-3}$ in.)	Beam current (μA)
Flat	2615	80	$5{\cdot}8 \times 10^4$	20 → 15	−2 → +10	80 → 30
	2680	40	$9{\cdot}6 \times 10^4$	15 → 20	0 → +10	100 → 50
	2760	16	$1{\cdot}8 \times 10^5$	10 → 15	2 → +10	150 → 50
	2840	8	$3{\cdot}1 \times 10^5$	5 → 10	5 → +10	200 → 80
22½° re-entrant	2615	80	$5{\cdot}8 \times 10^4$	~10	~ −10	~50
	2680	40	$9{\cdot}6 \times 10^4$	8 → 10	~ −10	~60
	2760	16	$1{\cdot}8 \times 10^5$	6 → 8	~ −10	~80
	2840	8	$3{\cdot}1 \times 10^5$	3 → 4	~10	~200

at 50 kV operating voltage. The results suggest a shield shape intermediate between the two investigated (10° re-entrant angle) as being a good compromise. With this,

any normally required total current can be obtained at any desired operating temperature and with optimum brightness and saturation.

In Conclusion

This chapter has been devoted exclusively to a description of the characteristics of what may be called the conventional type of electron microscope gun. The cold cathode gun of Induni and the field emission gun have been briefly mentioned. To these might be added the positive focus guns which have been described by Bruck and Bricka (1948) and by Steigerwald (1949). The latter guns are designed to give a focused spot at some distance below the anode, the advantage being that, in some applications, the condenser lens may be omitted. It is impossible to give any useful comparison between these various other types of gun and the conventional one since, although various claims have been made for them, they have not received the careful study which has been given to the conventional one.

However, since it has been shown that the conventional design of electron gun gives a performance approaching the theoretical, as far as its brightness goes, it is not to be expected that modification of the electrode system can give a higher picture intensity. The only way to achieve this would be through an improved cathode.

APPENDIX

Brightness

The brightness invariancy is an extremely important and useful one in electron optics. It was originally used by Langmuir (1937) who, however, limited its validity to the conjugate planes of a lens and the absence of aberrations. These limitations have been stated or implied in many subsequent papers and the principle, even with such limitations, wrongly applied.

The definition of brightness, or current density per unit solid angle, must refer to a vanishingly small area. That is,

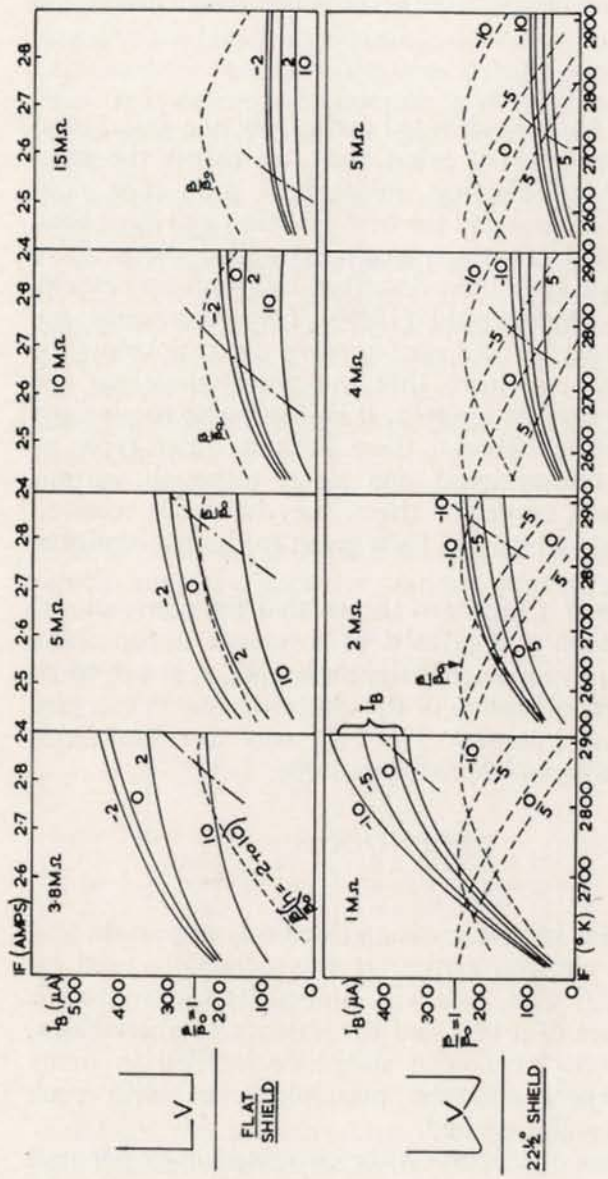


FIG. 6.11. *Variation of beam current and the ratio of brightness obtained to maximum brightness with heating current and cathode temperature for various values of cathode bias resistance and cathode height and three cathode shield geometries.*

the quantity refers to a point value. It is best thought of in terms of a very small aperture placed in the electron stream, at the point in question, perpendicular to the direction of electron flow. The brightness is then the transmitted current divided by the area and the solid angle of spread of the transmitted electrons. It is well to remember that serious error *can* result if the principle is applied to a large area, such as a whole focused spot, and an explanation for this has been given (Haine, 1957).

It may happen that the distribution of transmitted electron density is a complicated function of the angle. In such a case the whole matter becomes rather complex, though the principle is unaltered. There does not appear to have been either a rigid analysis or a simple description of the theory behind the principle of brightness invariancy in the general case, a much to be regretted omission. The only proof which has come to the present author's attention is that by D. Gabor in Myers' text-book *Electron Optics* (1939). The proof there given, unfortunately, contains several serious misprints.

The basic brightness relation says that, in absence of absorption, the current density per unit solid angle over all the space of an electron optical system is dependent only on the potential (relative to the emitting cathode) of the point of measurement. The value of brightness is given by:

$$\beta = \rho_c \phi_0 / \pi k T \quad (e\phi_0 \gg kT)$$

where ρ_c is the emission current density, ϕ_0 the space potential relative to that of the cathode, k is Bolzmann's constant ($8{\cdot}6 \times 10^{-5}$eV/°K) and T the absolute cathode temperature. This relationship can simply be proved in limited cases as, for example, by Langmuir, for the object and image and back focal plane of a lens. It can also easily be proved for a case where no imaging exists for points in a uniform accelerating field produced by parallel plates, where a limited area of one plate emits electrons. It is not always obvious that the brightness relation can hold generally. For example, when a lens is focusing a beam of electrons from a gun to a small spot, the addition of

spherical aberration to the lens will clearly increase the focused spot size and reduce its current density. At first sight the solid angle of the beam is unaffected and, hence, the brightness would appear to be reduced. Further consideration will show that the solid angle is, in fact, reduced, since any elemental area in the aberration caustic only receives electrons from part of the lens aperture instead of all of it. Therefore a small aperture placed in the beam will transmit a beam of angular cross-section less than the total lens aperture. If the aperture is moved radially the solid angle of the transmitted beam will reduce in proportion, as the current transmitted reduces, until both go to zero, leaving the brightness unchanged and equal to that which would be measured in the spot if no aberrations were present.

A fairly formal proof of the invariancy of brightness, based on the Gabor proof previously mentioned, is as follows:

Consider a cathode (Figure 6.12) in front of which is any arbitrary electrostatic field. At any point in this field

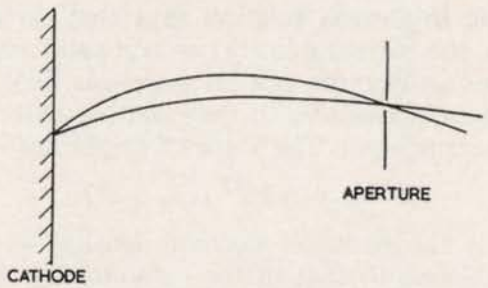


FIG. 6.12. *Diagram illustrating the derivation of the brightness invariancy relation.*

consider a very small aperture of area $\delta a'$. Now consider a small elemental beam passing through this aperture and spreading with a radial velocity $\delta v_r'$; i.e. with a divergence angle $\delta v_r/v'$, where v_z' is the velocity in the direction of motion of the elemental beam at the aperture. Suppose, then, electrons in this elemental beam leave the cathode from an area δa and with a spread of radial velocity δv_r.

The proof is based on Liouville's theorem which states that any 'bunch' of particles contained in an element of phase volume will continue to occupy the same element of phase volume as they travel, and no other particles can be present in this elemental volume. An elemental phase volume V is the product of an elemental space volume and an elemental velocity volume. Thus:

$$\begin{aligned} V &= \delta a.\delta z.(\delta v_r)^2\delta v_z. \\ &= \delta a.(\delta v_r)^2\delta v_z.v_z.\delta t. \end{aligned}$$

The density of electrons in the phase volume at the cathode is given by:

$$N(V) = \frac{\rho_c\,\delta a\,\delta t\,\delta\Omega/\pi}{e\delta a(\delta v_r)^2.\delta v_z.v_z\delta t}$$

where $\delta\Omega$ = solid angle of the elemental beam $= \pi\delta v_r^2/V_z^2$.

Hence:

$$N(V') = \rho_c/ev_z^3\delta v_z.$$

In the aperture:

$$\begin{aligned} N(V) &= \rho_1\,\delta a'\,\delta t/(e\delta a'(\delta v_r')^2\,\delta v_z'\,\delta t) \\ &= \rho_1/(e(\delta v_r')^2\,\delta v_z'.v_z'). \end{aligned}$$

By Liouville's theorem

$$N(V) = N(V').$$

Hence:

$$\rho_1/(\delta v_r')^2\,\delta v_z'.v_z' = \rho_c/v_z^3\,\delta v_z. \qquad (6.6)$$

Now:

$$\tfrac{1}{2}m(v_z)^2 = \tfrac{1}{2}m(v_z')^2 - e\phi$$

where ϕ is the potential difference between the point of measurement and the cathode.

Differentiating:

$$v_z\,\delta v_z = v_z'\,\delta v_z'.$$

Hence from equation (6.6):

$$\rho_1 = \rho_c(\delta v_r')^2/v_z^2.$$

The brightness

$$\begin{aligned}\beta &= \rho_1/\pi(\delta v_r')^2/(\delta v_z')^2 \\ &= \rho_c v_z'^2/\pi v_z^2 \\ &= \rho_c \times \tfrac{1}{2}mv_z'^2/\pi \times \tfrac{1}{2}mv_z^2.\end{aligned}$$

The energy $\frac{1}{2}mv_z^2$, may be put equal to the mean energy of emission of the electrons kT and hence:

$$\beta = \frac{\rho_c}{\pi}\left(\frac{e\phi}{kT}+1\right).$$

In most practical cases, $\phi/kT \gg 1$ and:

$$\underline{\beta = \rho_c e\phi/\pi kT.}$$

This is true of the elemental beam considered, but since no other electrons can occupy the same phase volume it will readily be seen that the value must hold for the total beam passing the aperture, since each elemental component will not only add to the transmitted current but also increase the solid angle in proportion.

The above proof is not quite as concise and formal as its originator would like, but has the merit of being easily understood. It will be seen that the result is quite independent of the distribution of the potential field and of the parts of the cathode which may contribute the elemental beams passing the aperture.

REFERENCES

BLOOMER, R. N. *Brit. J. Appl. Phys.*, **8** (1957), 83.
BRUCK, H. and BRICKA, M. *Ann. Radioelec.*, **3** (1948), 339.
DOLDER, K. T. and KLEMPERER, O. *J. of Electronics and Control*, **3**, 5 (1957), 439.
HAINE, M. E. and EINSTEIN, P. A. *Brit. J. Appl. Phys.*, **3** (1952), 40.
HAINE, M. E. *J. Brit. I.R.E.*, **17** No. 4 (1957), 211.
HAINE, M. E., EINSTEIN, P. A. and BORCHERDS, P. H. *Brit. J. Appl. Phys.*, **9** (1958), 482.
HIBI, T. *J. of Electron Microscopy* (Japan), **4** (Annual edition 1956), 10.

HILLIER, J. and ELLIS, S. G. *J. Appl. Phys.*, **20** (1949), 700.
INDUNI, G. *Helv. Phys. Acta*, **20** (1947), 463.
LANGMUIR, D. B. *Proc. I.R.E.*, **25** (1937), 977.
MYERS, O. *Electron Optics*, Chapman and Hall (1939).
STEIGERWALD, K. H. *Optik*, **5** (1949), 469.

CHAPTER VII

OBSERVATION AND RECORDING OF ELECTRON IMAGE

THE easy observation of the final image in the electron microscope is vital to the successful operation of the instrument. There are three important functions, depending on observation—the first is searching the specimen to find areas of most interest; the second is correcting astigmatism and other defects; the third is focusing. It can be said at once that, for reasons which will be described in succeeding paragraphs, it is never possible to see detail in the fluorescent image as clearly as it will appear on a photograph. Thus, very often, detail will be observed on a photograph which is not even detectable on the fluorescent screen. In view of this, it is clearly of the greatest importance to make the best possible use of the image, taking every possible step to aid the eye in observing it.

The intensity of the image is limited by the maximum current density which can be applied to the object. This is only sometimes limited by object heating, but is always limited by the practical and theoretical considerations in the electron gun which are discussed in Chapter VI.

For a given maximum value of current density at the object, the image intensity on the screen depends on the magnification (varying as $1/M^2$), and the efficiency of the phosphor in converting the electron energy to light. The effectiveness of the eye in 'seeing' fine detail in the image depends on the size of the detail, its shape, its contrast relative to its surroundings, the state of dark adaptation of the eye, the colour of the emitted light and its intensity. It is at once apparent that a complicated compromise is necessary and it would only confuse the issue to deal with all these factors in detail.

The visual acuity of the eye falls off with intensity and contrast. Thus an improvement would result from increasing the size of the observed detail, by increasing the electron optical magnification, except that this still further reduces intensity. Thus, there must be some optimum value of magnification giving best viewing. Let us consider what can be done to optimise the viewing conditions. Clearly, maximum efficiency of the electron-light conversion is a first step. In a search for optimum phosphors it has been found that efficiencies up to 100 lumens/watt can be obtained with various available phosphors (Bril and Klasons, 1952; Einstein, 1954). This represents an energy efficiency of about 25% and there is little hope of any very drastic improvement on this. The efficiency of a phosphor increases with electronic accelerating voltages up to 30 or 40 kV, remains roughly constant up to 100 kV, and then falls off, due to the fact that the electrons penetrate so deeply that the light is largely absorbed before being emitted from the free surface. Thus, an advantage is gained by working at higher voltages, at least up to 100 kV. The biggest improvement which might be made would be to modify the angular distribution of the emitted light so that more of it was directed into the narrow solid angle accepted by the eye. In normal direct observation a fraction of only about 2×10^{-4} of all the emitted light is collected by the eye. There seems only a small likelihood that a screen with directional emitting properties might be devised, but the possibility cannot be entirely ruled out. Improvement can, however, be obtained by using a wide-angle magnifying optical telescope to view the image. The reason for this is illustrated in Figure 7.1. In (a), the eye views the screen direct and collects light from a small area δA into a solid angle $\pi\alpha_e^2$, α_e being the semi-aperture angle of the eye ($\sim 0{\cdot}008$ rad). In (b), a magnifying lens is included of such an aperture as just to fill the aperture of the eye. It will now collect light into a solid angle from $\pi\alpha_0^2$ where $\alpha_0 = M\alpha_e$ (small angles). Thus M^2 times as much light is collected. However, this light is spread over an image M times larger and thus results in an image magnified M times but of the

same intensity as the original. The result arises from the diffuse nature of the screen emission. The extra magnification obtained in this way can be increased, in theory, up to the point where α_0 approaches $\pi/2$ radians; beyond this no further gain can result. Thus, the maximum magnification would be about 200 ×. In fact, owing to the cosine distribution of the emitted radiation, the intensity would be down by π times by the time this magnification was reached but, in any case, practical factors limit the useful value of optical magnification which can be used to between 10 × and 20 ×. Apart from the difficulty of making magnifying

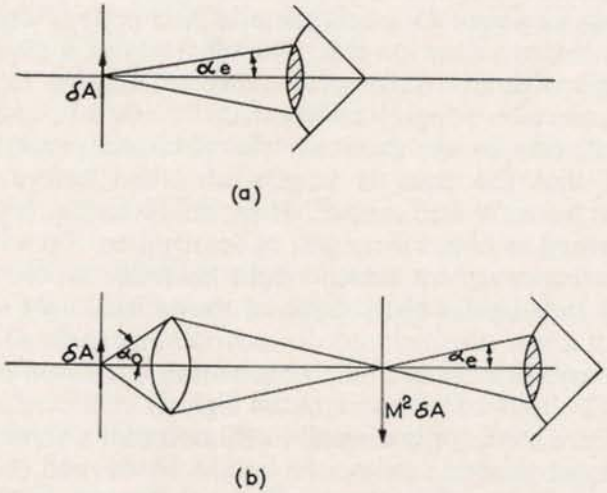


FIG. 7.1. *Diagram illustrating the advantage in using an optical magnifier.*

telescopes of very wide aperture, and of adequate working distance, there is no advantage in using greater magnifications since the resolving power of the fluorescent screen is limited, due to optical and electron scattering in its substance. The resolving power of powder screens has been investigated by Hinderer (1942) and von Borries (1948) who find values ranging from 70 microns at 50 kV to about 120 microns at 100 kV and 250 microns at 250 kV. In Figure 7.2 are shown two curves relating the visual acuity of the eye to the light intensity for detail of 20% and 40% contrast

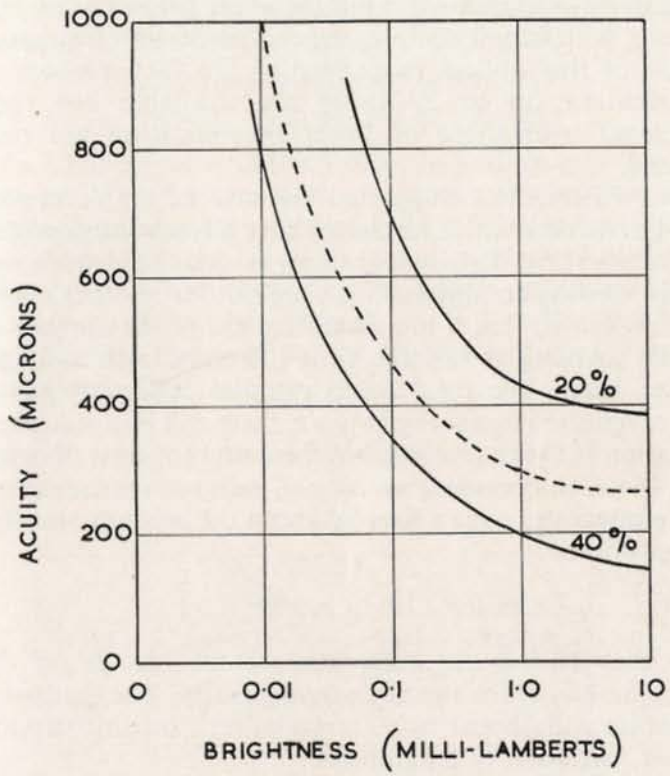


FIG. 7.2. *Variation of visual acuity with scene brightness (dotted line shows Agar's results for Fresnel fringes in the electron microscope image).*

respectively. The dashed lines show the variation in acuity measured by Agar (1957) for the observation of Fresnel fringes in the electron microscope. It is clear from these curves that the loss in acuity (for 40% contrast) increases very rapidly for an intensity below about 0·03 milli-lamberts, where the acuity is in the region of 400 microns. With a screen resolution of 80 microns it would appear that optical magnification of 5 times would be adequate. However, it is always of advantage to add, at least, some extra magnification if this can be done without loss of intensity. The difficulty of designing an optical

magnifier of adequate working distance, field of view and aperture, with magnification greater than 10 times, suggests the use of this optical magnification. In fact, viewers of magnification up to 20 times are available but their advantages over those of lower magnification are only marginal.

von Ardenne has suggested the use of single crystal phosphor screens which he claims have a resolving power of only 10 microns. This would allow a significant improvement in viewing conditions if it were not for another effect. It is well known that if in a measurement of the number of particles arriving at random time intervals, such as in an electron beam, the total mean number collected over a series of repeated measurements is n, there will be a statistical fluctuation between one measurement and the next of order $1/\sqrt{n}$. Now, in observing an area of radius r on the screen the eye integrates over a time of about 0·1 seconds and this number will be:

$$n = 6{\cdot}3 \times 10^{18} \times \rho_i \times \pi r^2 \times 0{\cdot}1$$

where $6{\cdot}3 \times 10^{18}$ is the reciprocal of the charge on the electron and ρ_i is the image current density. The statistical fluctuation will appear as a 'scintillation', or time-varying contrast, variation of magnitude:

$$1/\sqrt{n} = 7{\cdot}1 \times 10^{-10}\sqrt{\rho_i r^2}.$$

For a screen operating at 10^{-4} lamberts with an efficiency of 100 lumens/watt and an electron energy of 100 kV giving a current density of 10^{-11} amps/cm^2, the fluctuations over a 100-micron diameter area amount to about 3%. The increased fluctuation amplitude for smaller resolution areas would prohibit any advantage being taken of a higher resolution screen as suggested by von Ardenne.

In practice it is not possible to work at an intensity so close to the limit where detail is only just perceptible. Apart from the inconvenience of maintaining full dark adaptation, which means also a very low ambient light level, it is desirable to see contrast at least as low as 20%.

Experience shows that a screen intensity of at least 3×10^{-4} lamberts is desirable, or current densities of:

6×10^{-11} amps/cm^2 at 50 kV;
3×10^{-11} amps/cm^2 at 100 kV.

In Chapter VI is derived the relationship (equation 6.4) between screen current density, screen resolution and the required cathode current density (ρ_c). Also, in that chapter, it is shown that the life of the tungsten filament is given by $32/\rho_c$ hours. From these and the screen resolution figures, already given in this chapter, we can relate the screen brightness to filament life, for different voltages. Examples are shown in the following table.

TABLE 7.1

Cathode conditions for a screen intensity of 3×10^{-4} *lamberts*

	ρ_s	δ_s	ρ_c	t_l
(kV)	amps/cm^2	microns	amps/cm^2	hr
50	6×10^{-11}	70	1·3	25
75	4×10^{-11}	100	1·8	18
100	3×10^{-11}	120	2·0	16

It is seen that, for the required screen intensity level, the filament life is not long, and it must be emphasised that the above figures are based on optimum conditions. They assume a perfect electron gun and an illuminating aperture of $\frac{1}{3}$ of the critical value as is usually required to give adequate coherence. Suffice it to show that conditions are such as to make the viewing conditions impose a heavy strain on the electron gun design. It is fortunate that the eye, coupled with the brain behind it, can often improve upon what it sees by interpolation. The process is referred to elsewhere in this volume but is repeated for completeness here. In the important operations required to be carried out with the aid of the picture on the viewing screen, a process of adjustment is required to optimise, for example,

the focus, the fringe symmetry, *et al.* Since these involve setting a control to a datum position, on either side of which a deterioration of geometry of the picture structure occurs, it is possible by the interpolation process to obtain a setting more precise than is really observable. To do this the control is turned one way until a barely observable effect appears. It is then turned the contrary way, likewise until a barely observable effect appears. Next, the control is set half way between these extreme positions. If the effect is symmetrical a very good approximation to the correct setting is obtained. In astigmatism correction (see Chapter V) the effect is symmetrical and an improvement of five or even ten times can be obtained by the interpolation process. In focusing the effect is usually not symmetrical since phase effects tend to reverse edge contrast on either side of focus. Nevertheless, an improvement of several times can be obtained.

The Possibilities for Image Intensification

It is of special interest to consider the possibilities of intensifying the electron microscope image so that it can be viewed without dark adaptation or, alternately, so that the object loading and gun requirements can be reduced (Haine, 1955). Some methods of intensification also allow the possibility of contrast expansion, which would be most valuable. The basic problem involved is that of obtaining more light, in the visible image per electron, in the electron image, and it is first pertinent to consider to what extent limitation due to quantum effects will arise. It has already been shown that the quantum fluctuation limit is reached when a fluorescent screen is observed through a magnifying telescope and that, with the short exposure time available for the eye, these fluctuations prescribe a limitation which makes it impossible to see detail which can be recorded on a photographic plate with a longer exposure time. The use of an intensifier cannot improve this situation. Only higher illumination intensities or longer exposures can provide the necessary reduction of quantum fluctuations. If the former is ruled out then longer exposure times in visual observation

can still be considered. The use of a long persistence phosphor, or an equivalent time-integrating device in the intensifier, could provide a solution. However, the introduction of such a process would necessitate the processes of focusing, etc., being slowed down in proportion. It is probable that this would not prove unduly tedious when only applied in the most exacting work, but certainly an alternative display of low time constant would have to be provided for more routine work, for example, for use during searching of an object area. However, an increased intensity without increase of image resolution is still very much worthwhile, to improve the ease of viewing and eliminate the need for dark adaptation, etc.

There are a number of possible ways of achieving an intensification of the electron image, all involving energy conversion. These can be classified into two main types, static and scanned. A simple example of the first is the type of intensifier now being used for intensification of X-ray images (Coltman, 1948). This is illustrated in the diagram of Figure 7.3. The electron image is converted to light at

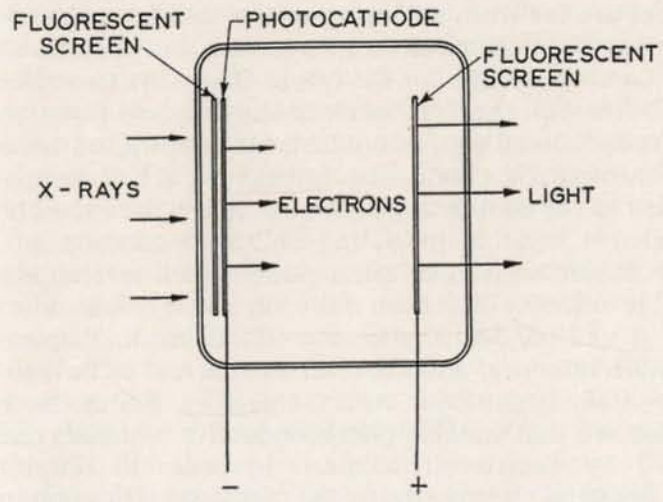


FIG. 7.3. *Diagram illustrating the static image intensifier.*

the phosphor screen A, in close proximity to a photo-emissive cathode B. The electron image emitted by B is accelerated and focused into the phosphor screen C, producing a light image. In the acceleration process energy is injected into the image and the final intensity may thereby be greater than that at the first phosphor screen. If the electron optical accelerating system demagnifies the image a greater intensity gain is obtained, a practice being followed in the X-ray intensifier.

The gain obtainable at unit magnification with existing photo-cathodes, and by the application of practical accelerating voltages, is likely to be too small to make the method worthwhile. In addition, practical difficulties preclude the possibility of placing the photo-cathode in the same vacuum as the electron microscope. This difficulty could be overcome by including an optical focusing system between the phosphor and the photo-cathode, but even the most efficient system so far devised would lose about 90% of the emitted light.

Further examples of the static type of intensifying system, which may well have great importance in the future, are the solid state intensifier screens. Up to the present these devices are far from achieving results which would render them useful in the electron microscope, but they do show considerable promise for the future. There are two relevant possibilities: in the first an electroluminescent layer and a photoconductor layer, in contact, are sandwiched between two transparent conducting layers. An a.c. potential is applied to the conducting layers but in the dark the photo-conductive layer is insulating and no excitation of the electroluminescent layer takes place. When an image of a suitable exciting radiation falls on the photoconductive layer it will conduct, by an amount depending upon the radiation intensity, and allow an a.c. current to flow to the electroluminescent layer which, therefore, fluoresces. It is well known that suitable photoconductive materials can be excited by electron irradiation. In order to obtain an intensification, apart from the necessity for obtaining adequate photosensitivity and adequate electroluminescent

efficiency, it is necessary that the photoconductive layer be matched to the characteristic of the electroluminescent layer. Successful light amplifiers have been constructed by these means but, so far, they have suffered from the disadvantage of having a very long time lag and a limited resolution (Kazan, 1957).

The second form of solid state intensifier, due to Cusano (1955), is possibly of more direct interest to the present application, and consists of a thin film of suitable phosphor sandwiched between two transparent conducting layers. The application of a suitable exciting radiation causes fluorescence which can be increased in intensity by the application of a d.c. potential across the conducting layers. A successful intensification has been achieved when exciting such layers by ultra-violet light, the emitted light being in the visible region. It is necessary that the phosphor layer shall be homogeneous and this result has been achieved by preparing the layer by evaporation or vapour condensation techniques. There has been no experimental evidence, so far, that such a screen could be excited by electron irradiation, but there is good reason to suppose that such would be the case.

Much more promise can be held out for the scanned type of system. This might consist of the direct application of a television camera tube to view the phosphor screen via an optical lens, but existing camera tubes are not sufficiently sensitive to give useful advantage. An alternative method, developed by Haine, Einstein and Ennos (1958) is to build into the microscope a camera-type tube responding directly to the electron beam. Figure 7.4 illustrates the system diagrammatically. The sensitive screen (S) consists of a thin film of amorphous selenium about 50 microns thick, supported on an aluminised film of the plastic material 'Mylar'. The film is mounted across a $2\frac{1}{2}$ in. aperture in a supporting plate. The aluminium backing (B) acts as the signal plate, as will be described. The front free surface of the selenium is scanned with an electron beam from a low-voltage electron gun in a television-type raster. The beam from the gun is accelerated to about 1000 volts and

then decelerated so as to arrive at the sensitive layer with low velocity. Deceleration is carried out in the field produced by the decelerating rings (C) which are adjusted at

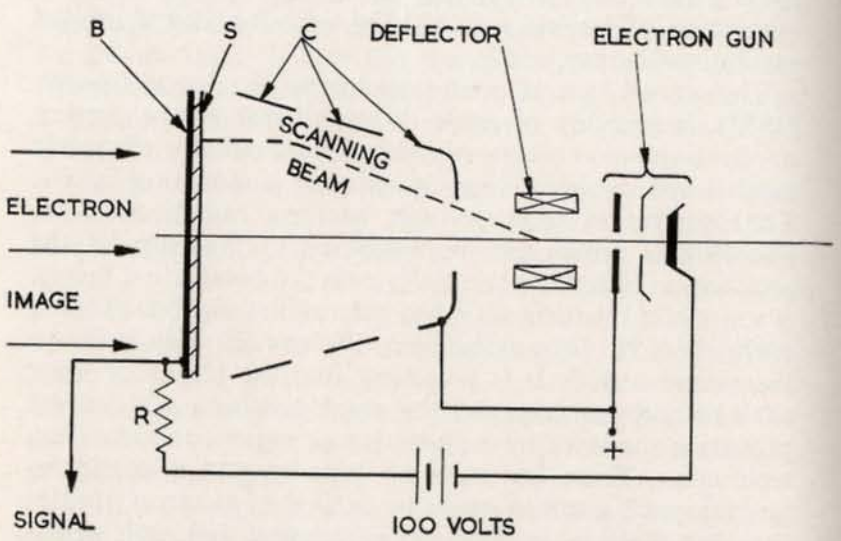


FIG. 7.4. *Diagram illustrating the scanned type image intensifier.*

such potentials as to cause the scanning beam to fall orthogonally on the selenium. The signal plate is held at a potential of 10 or 20 volts positive with respect to the gun cathode. The system is essentially similar to that used in the 'Vidicon' type television camera tube (Weimer, etc., 1951).

The scanning beam charges the surfaces of the screen negatively until it reaches cathode potential. Thereafter the scanning beam cannot reach the screen and is reflected away from it. This process is known as stabilising the potential of the screen, and there is now a potential difference of 10–20 volts across the selenium layer. During the frame period, between the moment when the electron beam crosses any particular element of the screen in one frame and the next, the charge stored on the screen will leak away through the film by an amount depending on the

resistivity of the selenium. This is very high in the absence of radiation, but when irradiated with high-energy electrons the resistivity falls due to the excitation of hole-electron pairs. The holes have high mobility and low probability of capture. The areas of highest excitation intensity will develop lowest resistivity. At the end of a frame the electron beam recrosses the discharged area and recharges it. The collected charge induces a like charge displacement in the signal plate, by capacitive action, and a signal potential proportional to this displacement is generated across the load resistance (R). This signal can be amplified and used to modulate a display cathode ray tube scanned in synchronism with the pick-up tube. Non-linearities may be introduced in the amplifier to improve contrast.

The basically important consideration deciding the possibility of this system is whether adequate conductivity can be induced in the selenium layer to give a useful signal under the conditions of interest in the electron microscope. Without analysing the operation in too much detail, it can be said that a discharge voltage of 0·2 volts is about the minimum which will give a useful signal, due to the thermal energy spread of the electrons in the scanning beam. The problem, therefore, is to ascertain whether 10^{-11} amps/cm^2 of excitation current will produce enough conduction current to discharge the screen voltage by 0·2 volts in one frame period. Spear (1956) has shown that with 50 kV electron excitation a current multiplication of about 2000 can be obtained. Thus a conduction current of 2×10^{-8} amps/cm^2 would result in an image current density of 10^{-11} amps/cm^2. The dielectric constant of amorphous selenium is 5·8 hence the capacity/unit area for a 100 micron thick film is about 5×10^{-11} mF. The charge dissipated in, say, 1/25 sec would be:

$$q = 2 \times 10^{-8} \text{ amps} \times 1/25 \text{ sec}$$
$$= 8 \times 10^{-10} \text{ coulombs.}$$

The discharge voltage then becomes $8 \times 10^{-10}/5 \times 10^{-11}$ ≈ 16 volts.

This should provide a very adequate signal and give a displayed image of equal quality to that on the fluorescent screen, but, of course, at a much higher intensity.

The resolution of the final displayed image is determined largely by the size of the scanning electron beam in the image tube. It is pertinent to consider whether by reducing this below the 80 microns resolution of the fluorescent screen a better image could be obtained on the display. The first limitation which will arise is electron noise, though a reduction by a factor of 3–4, thus bringing the resolution equal to that of the photographic plate, will only introduce about 10% contrast at 10^{-11} amps/cm^2 screen current density. This should still allow the system to be used with full success for reasonably contrasting images, such as in the Fresnel tests. By using a long afterglow display tube phosphor, so that several frames are effectively integrated, the noise could be reduced. For example, 0·5-second integration should make it invisible. Thus the system described offers the possibilities for producing an image as good as that which can be obtained on a photographic plate, and at full intensity for observation in a lighted room. Several displays could easily be run in parallel and non-linear amplifiers could be used to introduce controlled contrast expansion. Simultaneous observation and photography can be obtained and cinematography easily applied. An incidental advantage would be the easy application of electronic particle counting methods.

At the time of writing, this description of the apparatus has been made to present a picture which is almost comparable with that obtained on the fluorescent screen, but with a great increase in intensity. A reasonable picture can be obtained even at 10^{-12} amps/cm^2 in the primary electron beam when the fluorescent screen image is virtually invisible. There seems, at present, no reason to suppose that the theoretically predicted improvement, necessary to make the image comparable in resolution and contrast to that obtained on the photographic plate, will not be achieved in the fairly near future.

Photographic Recording of the Electron Image

It is possible to record, photographically, the electron microscope image, either by allowing the electrons to fall directly on the photographic plate or by photographing the fluorescent screen with a conventional optical camera. For a given electron intensity the exposure time required is not greatly different for the two methods if a wide aperture lens and very fast plate are used in the optical camera. However, the latter process introduces the limiting resolution of the fluorescent screen and is not normally used except for cine microscopy.

The two most important observations which can be made concerning photographic recording of the electron microscope image are, firstly, that the very wide range of speed and contrast found, when different photographic plates are exposed to light, no longer holds for electron irradiation. The difference in speed between, for example, photomechanical and high-speed emulsions, is less than 10 to 1 for electron irradiation. The second important observation is that the graininess of the recorded image with electron irradiation is almost entirely due to electron noise. A simple calculation shows that the degree of graininess in the electron exposed plate is just about what would be expected from considerations of the electron quantum noise. It follows from this that though some plates may be faster than others they will also show more quantum noise.

Another factor of interest is that many photographic plates exposed to electron irradiation give a photographic density which is directly proportional to the electronic charge falling on them (Digby, Firth and Hercock, 1953). The sensitivity increases with accelerating voltage up to a point where the electrons penetrate completely through the emulsion. After this the sensitivity falls off.

The fact that the graininess of the photographic plate is limited by quantum noise indicates that, as far as electron irradiation goes, the photographic plate is ideally efficient, and thus we can expect no significant improvement in the photographic process. The electron microscopist would

welcome a plate which, though still sensitive to electron irradiation, is insensitive to light. This would not only simplify the electron microscope camera design but also the subsequent processing of the plates. This desirable feature might be obtained by staining the emulsions with a suitable dye which renders them insensitive to light; the dye preferably bleaching out during the fixing process.

The perfection of an intensifying system for the electron microscope image may allow the photographic recording to be carried out from the intensified screen. The main advantage for so doing will lie in the simplification of the camera equipment, allowing the use of lower aperture optical lenses. Quantum limitations will still restrict the advantage of the shorter exposure times possible, for though it might be possible to obtain sufficient brightness to reduce exposure times to small fractions of a second, the photographs taken would be excessively 'grainy' due to quantum noise. It is again emphasised that this is a basic limitation due to the particulate nature of the electron beam in the microscope.

REFERENCES

AGAR, A. W. *Brit. J. Appl. Phys.*, **8** (1957), 410.
VON ARDENNE, M. *Z. Phys.*, **115** (1940), 339.
VON BORRIES, B. *Optik*, **3** (1948), 321.
BRIL, A. and KLASENS, H. A. *Phillips Res. Rep.*, 7 (1952), 401.
COLTMANN, W. J. *Radiology*, **51** (1948), 359.
CUSANO, D. A. *Phys. Rev.*, **98** (1955), 546.
DIGBY, N., FIRTH, K. and HERCOCK, R. J. *J. Photog. Soc. (England)*, **1** (1953), 194.
EINSTEIN, P. A. Unpublished.
HAINE, M. E. *Les Techniques Recentes en Mic. Elec. et Corpusc.*, Toulouse, 4th–8th April (1955). (Published by Centre Nationale de La Recherche Scientifique)
HAINE, M. E., EINSTEIN, P. A. and ENNOS, A. E. *J. Sci. Instrum.*, **35** (1958), 466.
HINDERER, K. *Z. Phys.*, **119** (1942), 347.
KAZAN, B. *Proc. I.R.E.*, **45** (1957), 1358.
SPEAR, W. E. *Proc. Phys. Soc.* (England), B.**19** (1956), 1139.
WEIMER, P. K., FORGUE, S. V. and GOODRICH, R. R. *R.C.A. Review*, **12** (1951), 306.

CHAPTER VIII

DESIGN CONSIDERATIONS

IT is clearly not possible to discuss all the manifold considerations which must be given to the question of detailed design of the electron microscope. From the user's point of view, the most important features are reliability, adequacy of performance and convenience in operation. This book is concerned more with providing a means of obtaining a reasonably easy understanding of the basic physical factors affecting performance than in describing the detailed engineering problems involved in instrument construction. Indeed, the attainment of the former must greatly reduce the difficulty of the latter. Thus, the following paragraphs will be devoted to a discussion on the more general matters affecting design and will not concern themselves with the many foibles and preferences of manufacturers and customers, which finally settle the overall layout and appearance of the finished instrument.

Examples of specific design features are taken, mainly, from the Metropolitan-Vickers Electrical Company instrument type E.M.6 (denoted M-V. E.M.6) (Haine and Page, 1956). Comparable features are incorporated in instruments manufactured in many parts of the world and may usually be found in the appropriate literature.

The electron optical system has, as its three special functions, to produce the necessary conditions for illumination, to provide an objective lens giving adequate resolving power and to provide a magnifying system giving the desired range of magnification without image distortion. It is desirable that the overall system shall be designed with economy in size and, of course, cost. Size is important as a more compact unit can give greater mechanical rigidity. The aim to achieve maximum simplicity in design is consistent with economy in cost and reliability in operation.

The Illuminating System

The main requirements for the illuminating system are to provide a beam of maximum brightness with the required divergence angle and illuminating spot size. The divergence angle must be adjustable between the maximum required value and a value of about one-tenth of this. The focused spot size is only important in that the total energy dissipated in a given specimen is proportional to the illuminated area. This is, of course, assuming a constant current density should be obtained since, for a given beam angle, the maximum current density obtainable is invariant, as is shown by the brightness relationship (Chapter VI). The temperature of the specimen is more dependent on the total power dissipation than the power density; therefore, to prevent specimen overheating, it is desirable to keep the illuminated area as small as possible. This is also important in order to minimise temperature changes in the specimen holder and stage, which may lead to thermal drift due to differential expansion effects.

The Electron Gun

The dependence of the performance characteristics on design and operating parameters of the electron gun has been fully discussed in Chapter VI. The other main factors affecting the design are the high-voltage insulation and the facilities for alignment. The former is a matter mainly of technical importance. The high-voltage insulator must be capable of operation *in vacuo* and this limits the choice of materials to glass, steatite or porcelain. The last is preferred. Modern microscopes are usually arranged so that the high-voltage cable is sealed into the gun insulator, thus avoiding the necessity for any air insulation. From an operating viewpoint it is worthy of mention that oil vapour, condensing on porcelain insulators, proves a serious cause of surface flashover. Oil vapour is, in any case, a nuisance in the microscope column and its avoidance is of great importance in minimising specimen contamination.

Alignment facilities should include a means of centering the filament in the cathode shield and, preferably, a means

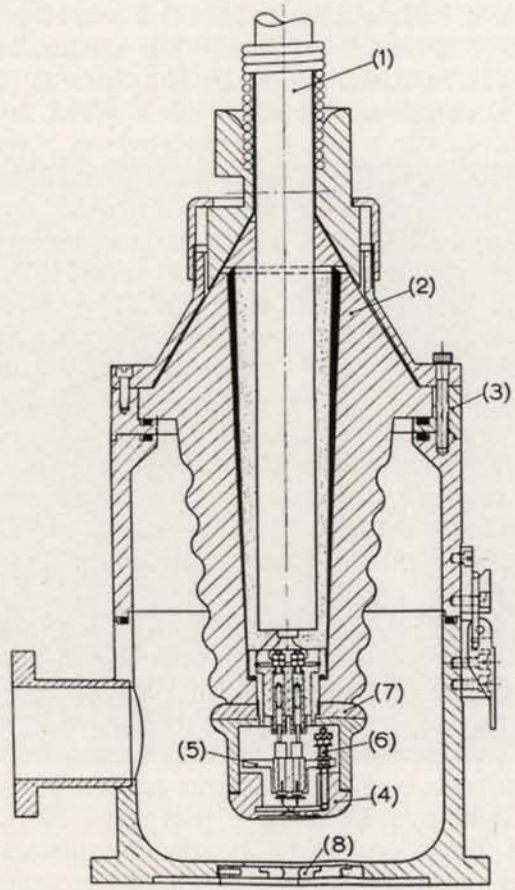


FIG. 8.1. *Cross-sectional diagram of M-V. E.M.6 electron gun.*

of centering the beam around the condenser lens aperture. In addition, the various parts of the gun should be accurately spigoted together to give an overall alignment. The centering of the filament is, in some instruments, achieved by accurate manufacture of the filament and its support so that, when plugged in, it is adequately centered. Other instruments provide adjusting screws for centering.

A typical example, the M-V. E.M.6 electron gun, is shown by the cross-sectional diagram in Figure 8.1. It

comprises the high-voltage cable (1) clamped and sealed into the porcelain insulator (2) which is vacuum sealed and clamped on to the main body (3). The cathode and shield assembly (4) comprises the shield on to which is mounted the cathode carrier plate (5) supported on three spring-loaded screws, one of which is shown (6); these screws allow the support plate to be raised, lowered or tilted to position the cathode tip. After so positioning a new filament the whole assembly plugs into the assembly mounted to the insulator end (7). A spare assembly can be kept with a new filament fitted and centered for quick interchange. The anode of this gun (8) consists of a plate sliding on a flat face in the bottom of the main body which can be centered by two perpendicular adjusting screws for beam alignment, as is described elsewhere.

To change the cathode assembly the top of the gun may be tilted back on a hinge, as shown in the photograph of Figure 8.2. This gun has a maximum operating voltage of 100 kV.

The true centering of the electron microscope filament is a matter of considerable importance. It can be seriously upset by distortion of the wire when first heated. For this reason filaments, as supplied by the maker, are usually pre-flashed in vacuum to ensure their subsequent stability. This operation causes the filaments to become somewhat brittle. It is of interest to note that most of the risk of subsequent distortion can be avoided by pre-heating to a dull red heat in air for a few seconds.

Condenser Lens

The function of the condenser lens is to converge the beam from the electron gun on to the object, and to control the convergence angle there. When the lens focuses the virtual source into the object plane, the convergence angle is that angle subtended by the condenser aperture. As the condenser lens power is varied the image of the virtual source moves axially either upwards or downwards; so long as the angle subtended by the source image at the aperture exceeds that subtended by the aperture at the

object, the illuminating angle is unchanged. When, however, further defocusing occurs the illuminating beam angle is reduced and is determined by that angle subtended by the virtual source image at the object (e.g. Haine, 1947).

FIG. 8.2. *Top portion of electron gun is hinged back to allow changing of cathode assembly at cathode.*

It has already been explained that the current density at the object, for a given beam angle, is prescribed by theoretical considerations. The total current falling on the object is determined by the total condenser lens aperture angle and the spot size. Thus, if heating is to be kept to a minimum and brightness maintained, the spot size must be made as small as possible; on the other hand the illuminating spot size cannot be made smaller than the total field of view which it is desired to observe. A reasonable compromise allows the virtual source size (as obtained from a 0·005 in. diameter filament) to be demagnified by about 15 times, thus reducing the minimum diameter of the

illuminating spot from about 15 microns to about 1 micron. To achieve this some modern microscopes are fitted with a 2-stage condenser lens system.

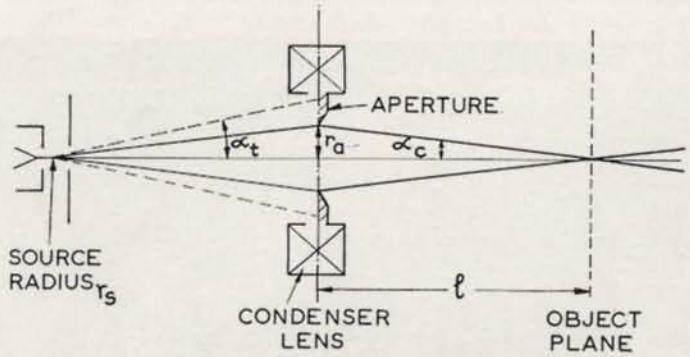


FIG. 8.3. *Diagram illustrating the single condenser lens system.*

Figure 8.3 shows diagrammatically the single-stage condenser lens system. The electron gun produces a cone of electrons of some maximum half-width angle α_t. The cone diverges from a virtual source of radius r_s which, as has been shown in Chapter VI, has a value of about 15 microns. The current density distribution across this source, as well as across the divergent beam, has a Gaussian shape. Near the optimum gun operating condition the current per unit solid angle will depend only on the accelerating voltage and cathode temperature, but the total beam angle will vary with the geometric shape of the gun electrodes. Thus, any total beam angle within wide limits may be obtained by a suitable choice of geometry. If, however, the optimum bias conditions are chosen then the current passing through an aperture in the beam, small compared with the total beam angle, will be constant. Since, also, the source size is little affected by the geometry, the focused spot intensity (current density) will be unaffected by these changes. To economise in total beam current the total beam angle should be kept small. It should, however, be large enough for the part transmitted by the condenser aperture to constitute only the central part of

the Gaussian distribution, otherwise a loss of intensity results. To achieve this, the half-width of the beam should exceed about twice the condenser aperture diameter.

A further consideration of importance, in choosing the total beam angle to be used, arises from the fact that a small misalignment of the filament in the shield aperture results in the emergent beam being tilted with respect to the gun axis. If this tilt is excessive the central part of the beam may miss the condenser aperture. The larger the total beam angle the wider can be the alignment tolerance on the filament centering. A greater tolerance can, of course, be allowed if a means for aligning the beam with the condenser aperture is provided. This alignment can be achieved by provision of controls to:

(a) align the filament in the shield,
(b) align the condenser aperture with the beam, or
(c) to align the gun with the aperture.

(a) is only practicable as a pre-set arrangement, allowing the filament to be centered on insertion, and requires great care in filament design and manufacture to avoid a small subsequent deflection on heating the filament: (b) is not very practicable as it results in the aperture being misaligned in the condenser lens: (c) is the step usually adopted. It can be achieved by the introduction of a transverse movement of the gun with respect to the condenser lens and aperture, but a much simpler arrangement can be achieved if the anode plate is given a lateral displacement in two dimensions, perpendicular to the axis. Since the aperture in the anode, in association with the accelerating field, constitutes a divergent electron lens, a lateral displacement results in a bending of the transmitted beam. The focal length of such an aperture lens is independent of its diameter which is, therefore, not at all critical. The adjustable anode plate has already been referred to in the previous section (see Figure 8.1).

The semi-angle of the focused beam converging on to the object has a maximum value determined by the radius of the condenser aperture and given simply by r_a/l, where

l is the aperture to object distance. It is usual to use an aperture radius which gives a maximum beam angle approximately equal to the optimum objective aperture angle for best resolving power (equation (3.3)). When the condenser lens is defocused more than a certain amount, the area of illumination increases and the beam angle is reduced, thus increasing the coherence of the illuminating beam. Since it is rare for the electron microscope to be used with the illumination angle as large as the optimum objective angle (critical illumination), it is permissible to use a condenser aperture smaller than that giving the full angle and thus to reduce somewhat the object loading. The same aim can be achieved by a field aperture immediately above the object but this must be very small (e.g. 50 microns) and presents difficulty in alignment. The use of a double condenser lens described later is a better method.

For reflexion electron microscopy (see Chapter IX), the image is produced only by low-angle scattered electrons. Even at the comparatively low magnification used a low image brightness is obtained, and the highest possible intensity in the illuminating spot is required. It is often, then, of advantage to increase the illuminating beam angle considerably beyond the value used for transmission working. An interchangeable condenser aperture is advantageous in this respect. The M-V. E.M.6 instrument incorporates three apertures, mounted on a transverse rod,

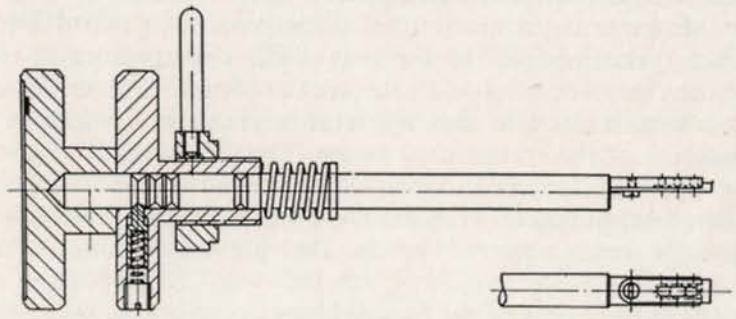


FIG. 8.4. *Photograph of interchangeable condenser aperture for the M-V. E.M.6 instrument.*

which enter the instrument radially just below the second condenser lens. The rod aperture carrier is shown in Figure 8.4.

The double condenser lens system is shown diagrammatically in Figure 8.5. The first lens (C_1) is strong and

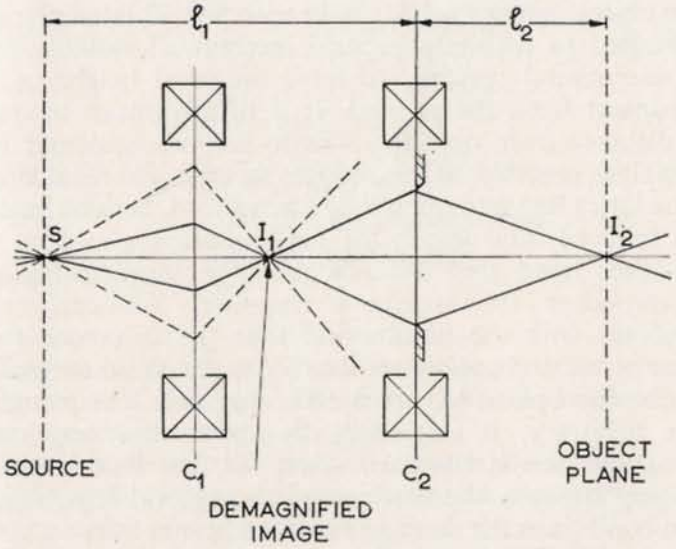


FIG. 8.5. *Diagram illustrating the focusing of the double condenser lens system.*

produces a demagnified image of the source (S) at (I_1). The second lens (C_2) must be included to project an image of the demagnified source on the object. The need for the second lens is prescribed simply by the necessity of providing a long axial 'working space' between condenser lens and objective lens. In the absence of this requirement, a single lens could just as well produce the demagnified spot. A further advantage in using two lenses is that the demagnification can be continuously varied, over a wide range, by variation of the power of the first lens. It will readily be appreciated that the use of a demagnifying lens system can never produce a higher intensity since the maximum usable beam angle is always limited by imaging

requirements and, coupled with the brightness invariance, this sets a limit to current density independent of condenser lens magnification.

It is of interest to consider the factors which determine the final design of a double condenser lens system. Usually some upper limit to the total length of the instrument tube, from objective lens to gun, can be specified. This length must be limited to maintain general mechanical stability and, for operational reasons, to limit the total height of the instrument from the ground. It is of advantage to make the distance from objective lens to second condenser lens as small as possible, as this reduces the required focal length of the latter lens and minimises aberrations, besides leading to a reduced total length for the illuminating system. On the other hand this distance must be large enough to accommodate the specimen chamber. We can start, therefore, with the assumption that the distances from source to second condenser lens (l_1) and second condenser lens to object plane (l_2) are fixed and known. The former is more arbitrary. It can easily be shown that maximum demagnification is obtained when the first lens is placed half-way between the source and the second lens. Under these conditions the demagnification is given by:

$$M = \tfrac{1}{4}l_1^2/f_1 l_2$$

where f_1 is the focal length of the first lens.

It was shown, in Chapter I, that the minimum focal length of a magnetic lens is obtained when $V_r/(NI)^2 = 0{\cdot}008$, giving a value of focal length of $0{\cdot}32(S+D)$. The values of S must be chosen so as to ensure that iron saturation does not occur. For an accelerating voltage of 100 kV ($V_r = 1{\cdot}1 \times 10^5$ volts), the ampere-turn excitation becomes 3700. Assuming a maximum allowable flux density ($B = 4\pi NI/10S$) of 18,000 gauss, the minimum value of S becomes about 0·26 cm. With an S/D ratio of two, the minimum focal length then becomes $0{\cdot}32(0{\cdot}26 + 0{\cdot}13)$ cm $\approx$ 0·125 cm.

As a typical example, if l_1 and l_2 are put equal to 15 cm, the maximum demagnification becomes about 30 ×.

There are several difficulties in making use of this optimum design. In the first place, the flux density specified is higher than it is desirable to use and, secondly, the bore diameter of the lens (0·13 cm) is very small for accurate machining. A more practical design might have the following parameter values:

$$\begin{aligned} S &= 0{\cdot}36 \text{ cm} \\ D &= 0{\cdot}6 \text{ cm} \\ S+D &= 0{\cdot}96 \text{ cm} \\ f/(S+D)_{\min} &= 0{\cdot}31 \\ f_{\min} &= 0{\cdot}3 \text{ cm} \\ V_r/(NI)^2 &= 0{\cdot}0075 \\ NI &= 4000 \text{ ampere-turns} \\ H_p &= 14{,}000 \text{ gauss.} \end{aligned}$$

For $l_1 = l_2 = 15$ cm

$$M_{\max} = 12{\cdot}5.$$

This value of demagnification is adequate for practical purposes and the lens dimensions are more reasonable from a manufacturing point of view.

The design of the second, comparatively weak lens appears, at first sight, a simple matter. However, it requires some careful consideration to avoid the introduction of astigmatism and spherical aberration. It will be seen that, for the example quoted, the focal length requires to be about 5 cm. It is undesirable to have to impose symmetry tolerances for this lens as high as those required for the objective lens. The maximum permissible value of the longitudinal astigmatism is given by $z_a < d_s/M\alpha_c$, where d_s is the source diameter. This value would give just detectable astigmatism in the smallest spot. A slightly greater figure might be allowed since the spot is seldom used focused, and the astigmatism soon becomes negligible when the spot is defocused. For the usual source diameter of about 30 microns, and with $M = 12{\cdot}5$, $\alpha_c = 0{\cdot}005$ rad, the limiting value of z_a becomes 480 microns.

If a symmetry tolerance (δ) of not less than 10 microns is assumed, the maximum value of the astigmatism ratio (z_a/δ) becomes about 50. Equation (1.12) then gives values

for the excitation parameter $[V_r/(NI)^2]$ and thence can be obtained those for $f/(S+D)$ as given below for three values of S/D:

S/D	$V_r/(NI)^2$	$f/(S+D)$
0·6	0·13	3·4
1·0	0·10	2·5
2·0	0·066	1·6

The size limitation on the pole pieces for this lens is seen to be the opposite to that for the first lens. The size now tends to be too great, leading to the iron taking up too much space, which might be occupied by windings, and giving too great an axial penetration length for the fringing field. To reduce the size, the smaller value of S/D is chosen. Thus, for $S/D=0{\cdot}6$, imposing the above limitation to minimise the tolerance stringency, the following lens emerges:

$$
\begin{aligned}
f &= 5 \text{ cm} \\
f/(S+D) &= 3{\cdot}4 \\
S+D &= 1{\cdot}6 \\
S/D &= 0{\cdot}6 \\
S &= 0{\cdot}6 \text{ cm} \\
D &= 1{\cdot}0 \text{ cm} \\
V_r/(NI)^2 &= 0{\cdot}13 \\
NI &= 920 \text{ ampere-turns} \quad \text{at 100 kV} \\
H_p &= 1900 \text{ gauss.}
\end{aligned}
$$

The cross-sectional diagram in Figure 8.6 shows a typical design of double condenser lens. In this design the iron circuit is made in three pieces. The vacuum joints are made either side of spacers between the pole faces. The lens is demountable so that the coils can be wound on separate bobbins. The design of the magnetic circuit of lenses is discussed in a later section.

The Objective Lens

The main requirements for the objective lens are that it shall have small spherical aberration, astigmatism and chromatic aberration; it shall allow convenient placing

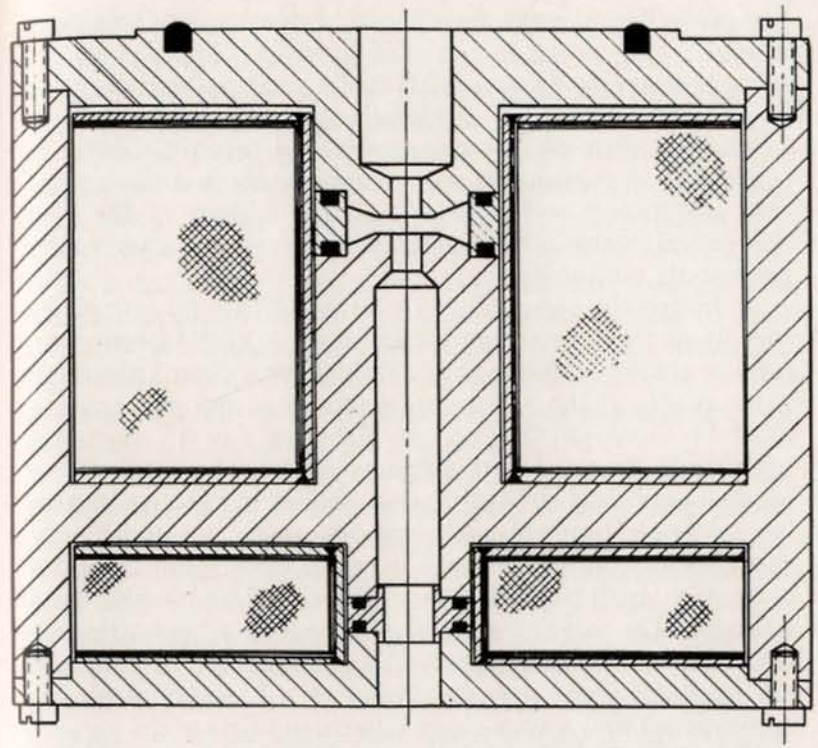


FIG. 8.6. *Cross-sectional diagram of double condenser lens from M-V. E.M.6 instrument.*

and movement of the specimen; it shall allow the fitting of a suitable aperture system and shall suffer from no defects which will upset alignment.

The requirement for low spherical aberration is not difficult to meet unless the very highest resolving power is required. Minimum spherical aberration is obtained with the strongest possible lens, but little is lost by quite a large reduction in lens strength (Chapter II). Chromatic aberration is also a minimum for the strongest possible lens, but much of the deleterious effect of chromatic aberration can be eliminated by adequate voltage and current stabilisation. Astigmatism can be corrected but correction becomes difficult if too large an amplitude is present in the

first place. Good stage stability and large areas of specimen available for searching are of practical importance. A reliable aperture system with easily replaceable apertures is also important. Consideration of all these factors inevitably leads to a compromise but, even so, little is really lost. A reasonable compromise leads to a theoretical resolving power well within a factor of two of the best theoretical value, but practical limitations can easily prevent its attainment.

Although the value of $f/(S+D)$ or $C_c/(S+D)$ falls fairly rapidly as the lens strength increases ($V_r/(NI)^2$ decreases), it must always be borne in mind that, for a given maximum value of H_p, if the excitation is increased the gap spacing must be increased. Hence, the value of f or C_c does not necessarily decrease with stronger lenses where a limit to the magnetic field strength is imposed, as in practice it must be. Le Poole (unpublished) has designed a lens with very short focal length and low chromatic aberration by using extremely small pole piece bore diameters and a high field strength. The advantage gained is mainly a less stringent electrical stability requirement; on the other hand, the arrangement has the disadvantage of mechanical inconvenience and a very close symmetry tolerance.

There are three possible ways of introducing the object holder or objective aperture into the objective lens: from above the lens, from below the lens or from the side—through the iron circuit into the magnetic gap. Since both object and aperture are necessary, and it is not really practicable to put both in by the same route, the object can be inserted into the top bore and the aperture in the side or bottom, or the object can be inserted in the side and the aperture from the bottom. These three possible combinations have all been used. The lateral insertion of the object is a simple and relatively cheap method and can be accomplished by the use of a simple rod-type airlock. The method suffers from the disadvantage that it leads to an inherently less stable arrangement mechanically. Also the necessity, in a high performance instrument, for a centrable aperture, which can be rapidly extracted for

cleaning, makes it desirable to use the lateral rod system for the aperture, thus requiring the object to be inserted from the top of the lens. A further argument in favour of such an arrangement results from the necessity of including an astigmatism corrector which can then be inserted through the bottom bore of the lens, or can be positioned in the bottom bore, since it can be positioned some way below the lens.

Having reached the above conclusions it is a small step to set a minimum practical pole bore diameter and spacing. Modern applications make it desirable to be able to scan a field in the object of at least 0·2 cm. To do this requires a pole piece bore of about 0·5 cm. The insertion of the aperture laterally into the lens gap requires a gap spacing of at least 0·3 cm giving an S/D ratio of 0·6.

It is of importance to consider how much may be lost by prescribing the vital dimensions of the pole pieces from these purely mechanical considerations. Figure 3.3 (Chapter III) shows that the resolving power as limited by spherical aberration and aperture diffraction is very little dependent on the bore diameter, provided a spacing near the optimum is used. The pole piece spacing should be kept small for a given flux density to conserve excitation. On the other hand it must not be so small that the resolving power falls off badly. Reasonable spacings, which give resolving power little worse than the optimum, are shown in the following table for 100 kV operating voltage and a range of flux densities. Corresponding values of excitation and resolving power are also given.

TABLE 8.1

Variation of optimum resolving power as a function of flux density

Flux Density (Kilogauss)	S (cm)	NI Amp-turns	d (Å)
3	1·6	3840	3·2
6	0·8	3840	2·9
12	0·4	3840	2·5
18	0·27	3840	2·2

This table shows that the optimum resolving power is little dependent on the flux density and that the minima on the d/S curves fall along a line of constant excitation. It also shows that a choice of $S = 0{\cdot}3$ cm requires a rather high flux density for an objective lens where such things as asymmetric stray fields must be kept to a minimum. A more reasonable choice would be $S = 0{\cdot}6$ cm, requiring 8 kilogauss and still 3840 amp-turns, to give an optimum resolving power of about 2·7 Å. The additional space provided is welcome. The main disadvantage of a large-bore diameter is that the diameter of the core of the iron circuit must be increased in proportion.

These conclusions are little different for 50 kV operation though slightly smaller spacing and excitation and lower resolving power result.

Having selected the main pole piece dimensions and the excitation, the rest of the iron circuit dimensions can be derived from the considerations described in the following section.

Magnetic Design

Experience has shown that an excitation coil not cooled with water can be operated at a loading of about 800 ampere-turns per square inch of cross-section. It can easily be seen that this loading figure is substantially independent of the wire size used, so that this can be chosen to give the most convenient operating voltage. In a water-cooled lens the excitation density can be increased to 1000 or 1200 ampere-turns per square inch depending on the effectiveness of the water cooling. It may be mentioned here that water cooling is advantageously applied to an objective lens to reduce temperature changes on the mechanical stage, which can be a serious cause of object movement leading to loss of resolving power. More will be said about such matters later in this chapter.

It has been seen that the maximum flux density in the lens gap does not very significantly affect resolving power. However, it does limit the minimum obtainable focal length and is of especial importance in projector lenses and, for

example, the first lens of the double condenser system. It has often been suggested that special iron alloys, such as the cobalt-iron combination, offer big advantages in this respect. Before one can assess the advantage of such alloys some criterion is necessary. Liebmann (1953) showed that little advantage is obtained in taking the gap flux density above a value where the permeability of the pole tips falls below 50. Applying this as a criterion it is easy to show that little or no advantage is gained in using cobalt-iron alloys. Furthermore, the published properties of these alloys usually refer to thin sheets, and depend on a heat treatment in hydrogen which is only partially effective with thick specimens. Thus, one may conclude that a fairly pure annealed iron is still the best material for pole pieces.

The optimum design for the iron circuit has been investigated by Mulvey (1953). The basic problem is how to shape the iron so that, with reasonable economy in overall size and winding space, the excitation loss in the iron is insignificant. This, broadly speaking, means keeping the flux density in the iron smaller than some upper limit which can be termed the saturation value, since, for optimum optical performance (lowest spherical aberration), the flux density in the gap, and hence in the iron in the immediate vicinity of the gap, must be a maximum, and this is limited to the saturation value. The requirement is to design the iron circuit so that the flux density is everywhere else small compared with the value in the gap.

A very simple design of lens which might be considered is shown in Figure 8.7 (a) as a cross-sectional diagram symmetric about the axis AA′. The gap is at B and the dotted lines show the flux lines. It will immediately be noted that, as one proceeds down the core away from the gap, more and more flux lines enter the iron. Therefore, since the cross-sectional area is constant, the flux density in the iron will increase to a maximum value at the end of the core opposite the gap (the root). The curves shown in Figure 8.7 (b) show the way the ratio of the flux density (B) in the core to the flux density b_0 in the gap increases along the axis, for three ratios of the gap spacing (S), to the core

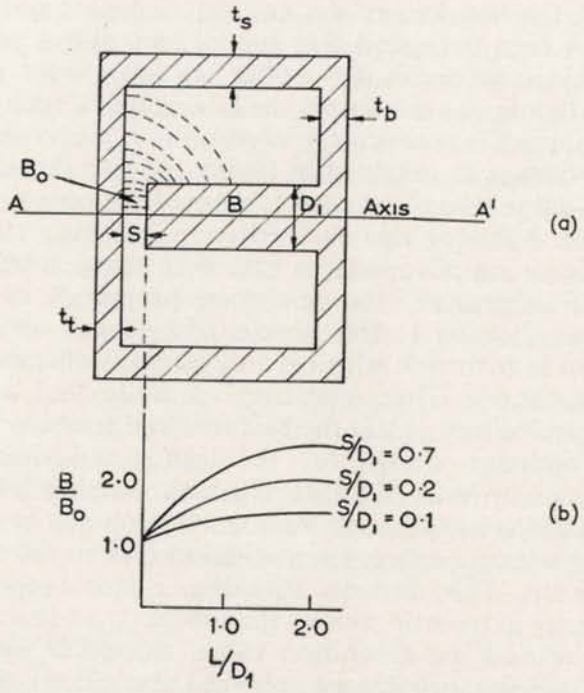


FIG. 8.7. (a) *Diagram illustrating unshaped pole piece design.* (b) *Relative flux density distribution in the core of the design shown in* (a).

diameter (D_1). Such a design clearly does not meet the requirements.

In Figure 8.8 (a) is shown a more practical design. The top of the core is here tapered so that, as more of the leakage flux is collected, the cross-section increases and the flux density can be made to decrease. As soon as the taper is discontinued the flux density begins to rise again, and Mulvey's investigations were aimed at ascertaining how far it was necessary to take the taper so that the flux density dropped far enough over the tapered region so as to ensure the subsequent rise never caused the density at the root to exceed the value in the gap. He found the angle of taper (θ) to have little effect between 45° and 70°, and the necessary condition to be achieved when the ratio of the core diameter

to the face diameter (D_2/D_1) to be at least 3 (see Figure 8.8 (b)). A further increase of diameter is necessary if the centre of the core is bored out to give an equivalent cross-sectional area of the solid core. For safety, when very high flux densities are used in the gap, a ratio of about 4 might be used.

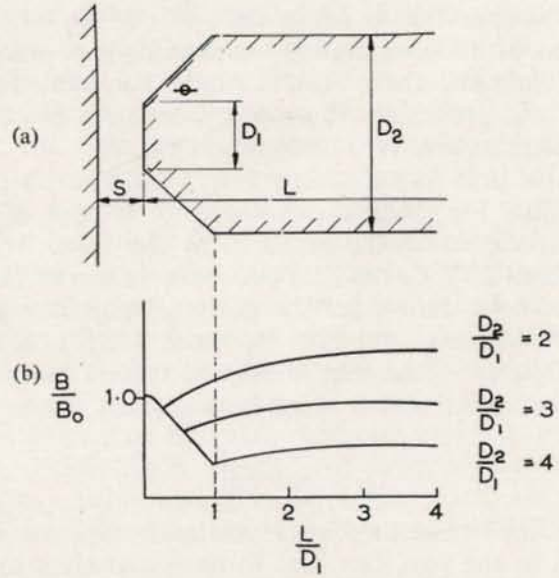


FIG. 8.8. *Tapered pole piece with resulting relative flux density distribution.*

Mulvey has also shown that the thicknesses of the bottom plate (t_t, t_b, Figure 8.7), outer shell (t_s), and top plate (t_t), which give maximum flux densities in these regions equal to the gap flux densities, are given by:

$$t_b \approx 0{\cdot}25D_2 \tag{8.1}$$
$$t_s \approx 0{\cdot}17D_2(S/D_1)^{1/3} \tag{8.2}$$
$$t_t \approx 0{\cdot}5D_1(S/D_1)^{1/3}. \tag{8.3}$$

To reduce the magnetic losses in these regions to a reasonable level, the shell thickness should be up to twice the value given by (8.2), while the values of t_t and t_b should be increased by at least 50%.

For a symmetrical lens, where the gap is in the middle of the core instead of at one end, the above values hold if S is put equal to half the pole spacing, and equation (8.1) is used for top and bottom plates.

The diameter of the top of the pole piece (D_1) should clearly be sufficient to allow the fringing effects due to the bore and the edge not to extend throughout the intervening radius. This condition is met if the value of $D_1 = D + 2S$ is maintained, although a slightly lower value is permissible.

As an example, these results might be applied to the objective lens previously discussed, where $S = D = 0{\cdot}6$ cm, $H_p = 8$ kilogauss and $NI = 3840$ ampere-turns.

A suitable pole face diameter (D_1) would be 1·5 cm. To accommodate the astigmatism corrector a bore of about 2 cm diameter would be required at the root. This will remove about 20% of the iron and increase the flux density correspondingly. However, the designed gap flux density is only 8 kilogauss, and even with the 2 cm bore and a ratio of $D_2/D_1 = 3$, the flux density at the root would not exceed about 10 kilogauss which is adequately conservative.

Thus:

$$D_2 = 4{\cdot}5 \text{ cm.}$$

The flux density in the bottom plate can be permitted to equal that in the gap, but that in the outer shell and top plate are preferably kept as low as possible, say, to half the gap value.

Applying this and equations (8.1), (8.2) and (8.3) we get:

$$t_b = 1{\cdot}1 \text{ cm}$$
$$t_s = 1{\cdot}0 \text{ cm}$$
$$t_t = 1{\cdot}0 \text{ cm.}$$

The flux density in the top plate falls off towards the centre and it is permissible to cut away the top face in a 60° conical hole at the centre. This helps to accommodate the object holder and mechanical stage. The excitation required is 3840 ampere-turns. With water cooling this requires a winding area of about 4 in^2. Thus, all the main dimensional parameters of the lens are determined.

Practical Lens Design

Much could be written on the subject of practical lens design. The requirements, especially for the objective lens, are rigorous. The high degree of symmetry required calls for great accuracy in machining. Good alignment between pole pieces is desirable if the overall instrument aligning difficulties are to be minimised. The magnetic material must have a high saturation flux density, and any stray magnetic fields must be at least symmetrical about the axis.

It has, for some years, been thought desirable to manufacture the pole pieces separately to the main magnetic circuit or yoke. In this way the pieces of iron on which the most accurate machining is necessary are kept small. On the other hand the use of separate pole pieces leads to the necessity of providing very accurately fitting surfaces between pole pieces and the lens body, and also leads to the danger of stray fields arising from these magnetic joints. The present author has demonstrated that lenses, at least as good in performance and simpler and cheaper in manufacture, can be made without separate pole pieces (Haine, 1954). Such lenses have proved particularly free from stray fields and give extremely good alignment. A typical example, illustrating the essential simplicity of the design, is shown in Figure 8.9.

The lens, which is fitted with water cooling and objective aperture system, comprises the two soft iron pieces (1) and (2) spigoted and bolted together on to the non-magnetic space piece (3) with vacuum seals at (4) and (5). The winding (6) is outside the vacuum and water cooled in the annulus (7). Through the spacer the aperture is inserted, via a rod-type air lock, described later. The whole lens can be dismantled and reassembled, the coil being wound on a separate bobbin, which can thus be wound separate from the iron circuit.

In order to ensure maximum accuracy of alignment between the different parts of such a lens it is desirable to machine to a carefully thought-out schedule. Such a schedule may well involve assembling the lens part by part

on the lathe before final machining of successive parts. In this way all important surfaces, i.e. bores and spigots, can be machined to a common axis. The major source of error then resulting arises from the 'slip', or regular axial movement of the lathe headstock with rotation. Adequate care may reduce the latter to a small fraction of a thousandth of an inch. Greater accuracy may be sought by the

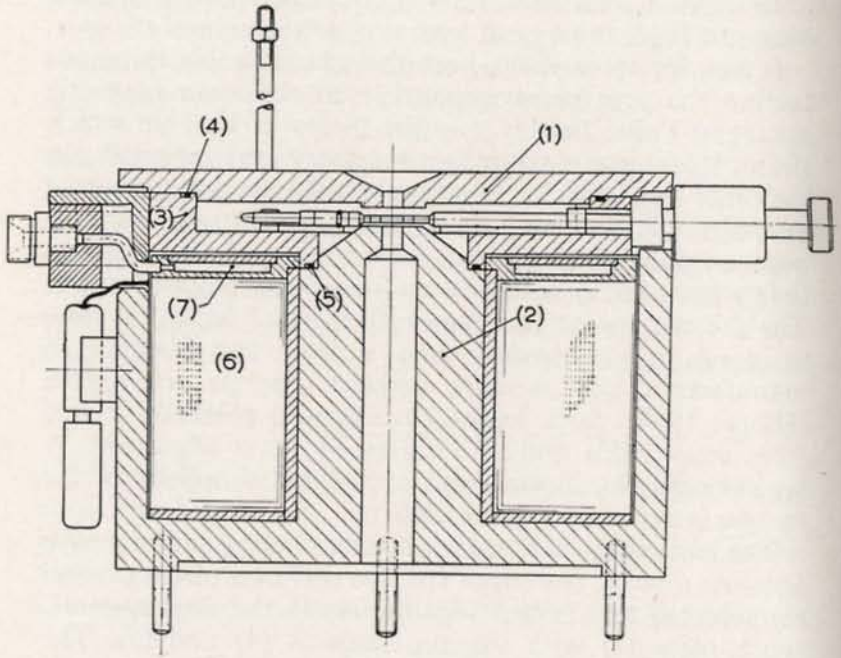


FIG. 8.9. *Cross-sectional diagram of objective lens (M-V. E.M.6).*

use of grinding methods but, in the author's experience, little or no advantage has resulted.

By the use of such techniques, not only for objective but also for condenser and double projector lenses, adequate mechanical accuracy can be obtained except for the small residual lack of axial symmetry which leads to astigmatism. Even this error can be regularly and reliably brought to a

sufficiently low value where correction is comparatively easy. It is highly probable that refined lapping techniques may make it possible to reduce asymmetry error almost to the extent where correction is unnecessary. Attempts have been described to achieve better symmetry by lapping, with little real success, but it is likely that the lenses on which these experiments were carried out had asymmetries in parts of the lens other than the pole pieces, and it may well be worthwhile pursuing this method with the modern type of lens, where it is known that only the pole pieces are likely to provide a limitation.

The Projector Lenses

The design and manufacture of the projector lenses can follow the same principles described for the objective lens. Symmetry tolerances become much less important for the final lens, and are mainly of importance in the intermediate lens when the instrument is used for small area electron diffraction. The basic design now depends largely on questions of focal length and, hence, magnification and also on distortion. Minimum focal lengths are often required, particularly for the final lens. This requirement leads to the use of the highest possible flux densities and especial care is necessary to provide an adequate iron circuit.

The requirement for a maximum magnification, up to, say, 500,000×, can, on its own, easily be met in a three-stage system. On the other hand, for a general purpose instrument, it is desirable to include means for varying the magnification over as wide a range as possible. Such a wide variation has been achieved in the 'Elmiskop' (Ruska, 1954) by a mechanism for interchanging the final projector lens pole pieces, which are mounted on a turret head rotatable from outside the vacuum, while the instrument is operating. Almost as wide a range of magnification can be obtained by the use of a three-stage magnifying system (objective plus two projector lenses) of optimum design. Since the analysis of this problem has not appeared in the literature up to the present, it is given here.

The range of magnification obtainable on a three-stage

instrument of given overall length is limited by distortion introduced by the projector lenses, on the one hand, and by the shortness of focal lengths, on the other.

In Figure 8.10, O is the objective lens, P_1 the intermediate lens, P_2 the final lens, and S the screen, OS being the instrument axis. $OP_1 = L_1$, $P_1P_2 = L_2$, $P_2S = L_3$.

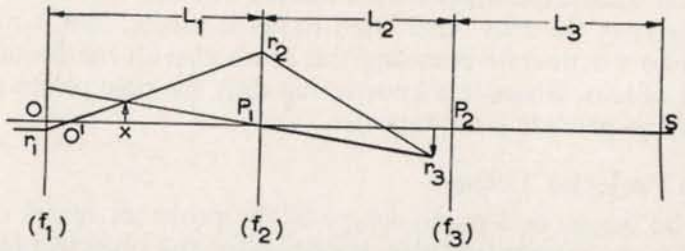


FIG. 8.10. *Diagram illustrating the derivation of the magnification range in a three-stage instrument.*

The maximum magnification is given by:

$$M_{\max} = \frac{L_1}{f_1}.\frac{L_2}{f_2}.\frac{L_3}{f_3} \qquad (8.4)$$

where f_1 is the objective focal length, and f_2 and f_3 are the minimum focal lengths of the intermediate and final projector lenses respectively.

The distortion of the final projector lens can be defined, in terms of the distortion constant (C_d, Chapter I), by the equation:

$$x = 4C_d\left(\frac{r_3}{D_3}\right)^2 \qquad (8.5)$$

where D_3 is the diameter of the final lens pole piece, r_3 is given by: r_s/M_3 or $r_s f_3/L_3$, where r_s is the maximum radius of the final image. Hence

$$x = 4C_d\left(\frac{r_s f_3}{D_3 L_3}\right)^2.$$

This equation defines a minimum length of L_3, below which the distortion exceeds the specified value x_0.

$$L_3 \geqslant \frac{r_s f_3}{D_3}\left(\frac{4C_d}{x_0}\right)^{1/2}.$$

This can be written:

$$L_3 \geqslant r_s.\left(\frac{f}{S+D}\right)\left(\frac{4C_d}{x_0}\right)^{1/2}\left(\frac{1+S}{D}\right) \qquad (8.6)$$

where S and D are written instead of S_3 and D_3, S being the lens pole piece spacing.

In Figure 1.12, $C_d^{1/2}(1+S/D)$ is plotted against $V_r/(NI)^2$. Also shown is $C_{sp}^{1/2}(1+S/D)$, C_{sp} being the spiral distortion constant. These curves are plotted from Liebmann's data (1952). It is seen that $C_d^{1/2}(1+S/D)$ varies little with lens power, rising slightly at the higher excitation and then falling sharply at excitations in excess of that required for minimum focal length. Unfortunately no advantage can be obtained from the low radial distortion of high excitation since the spiral distortion rapidly increases at high excitation, exceeding the radial distortion quite near the excitation for minimum focal length. The product of $f/(S+D)$ and $C_d^{1/2}(1+S/D)$ contained in equation (8.6), and hence the value of L_3, falls to a flat minimum at an excitation roughly corresponding to a minimum focal length. This gives an absolutely minimum allowable value of L_3 for any prescribed value of the distortion x. A small gain results from the use of a pole piece spacing to diameter ratio $(S/D) = 2$, the use of which ratio also gives a shorter focal length. Little gain is obtained below a value of V_r in $V_r/(NI)^2$ of 0·0075 or $f/(S+D) = 0{\cdot}30$.

Example:

$$\begin{aligned} V_r &= 1{\cdot}1 \times 10^5 \quad (100\text{ kV}) \\ x &= 0{\cdot}02 \quad (2\%) \\ r_s &= 5\text{ cm} \\ S/D &= 2 \end{aligned}$$

gives:

$$L_3 \geqslant 26\text{ cm}, \quad \text{and} \quad NI = 3820\text{ ampere-turns.}^*$$

It will be noted that, at this value of L_3, no range in focal length (or magnification) is possible without the

* It should be noted that very little change results if V_r in $V_r/(NI)^2 = 0{\cdot}01$ is used giving $NI = 3300$ ampere-turns provided the pole dimensions are scaled down to give the same flux density in the gap.

increase of distortion. The focal length can be chosen at any value above the absolute minimum value, as limited by pole piece saturation, by suitably scaling the size. For a saturation flux density of 18,000 gauss, the minimum spacing is 0·27 cm; the minimum focal length is 0·12 cm; the magnification is now 220. If a longer focal length is desired, it must be obtained by a direct scaling of $(S + D)$ in proportion to f, keeping the excitation constant, otherwise distortion will increase.

It will be noted that some useful magnification variation can be obtained from the second projector for search purposes if 10–15% distortion is accepted.

Intermediate Lens: Since all the useful magnification range must be obtained by variation of the strength of the intermediate lens, it is clear that, for the widest possible range, this lens must operate up to the highest magnification it can give, i.e. it must be designed to operate at its minimum focal length at the top limit of the magnification range. As the intermediate focal length is increased, the intermediate image moves axially from just in front of the intermediate lens towards the objective lens. Thus, the magnification of both the intermediate and objective lenses are reduced. At some point, distortion will eventually rise to a prohibitive value. Suppose this occurs when the intermediate image between the objective lens and intermediate lens is at the axial position X (Figure 8.10). Since the object is illuminated by electrons substantially parallel to the axis, all the electron rays leaving the objective lens pass through the back focal point of this lens. This point is depicted as O′ in the figure. Let the maximum radius of the first projector lens which can be used without distortion be r_2 (defined in position and radius by this symbol, see Figure 8.10). This radius corresponds to a ray leaving the object at a distance r_1 from the axis where:

$$r_1 = r_2 \frac{f_1}{L_1} \tag{8.7}$$

This immediately defines the minimum magnification since a field larger than r_1 cannot be imaged without exceeding

the specified distortion. Hence, the minimum magnification as limited by the distortion due to the intermediate lens is given by:

$$M_{\min} = \frac{r_s}{r_1} = \frac{r_s L_1}{r_2 f_1} \tag{8.8}$$

where r_s is the maximum radius of the final image, and other conditions are arranged so that the second intermediate image has a radius r_3, the maximum which can be imaged by the final lens. It now remains to relate the radius r_2 to the actual distortion produced.

The distortion in a thin electron lens is due only to the spherical aberration. This results in a ray passing through the lens, at a radius r_2, being reflected through an angle in excess of that required for correct focusing by an amount (δ) given by:

$$\delta = C_s r_2^3 / f_{21}^4 \tag{8.9}$$

where f_{21} is the focal length of the lens under the condition of minimum magnification.

The actual distortion in the second intermediate image measured as a fraction of its radius is given by:

$$\frac{\delta r_3}{r_3} = x = \frac{L_2}{r_3}\delta = \frac{L_2}{r_3}\frac{C_s r_2^3}{f_{21}^4}$$

corresponding to the extreme ray through r_2.

Hence, for a distortion x:

$$r_2 = \left[\frac{x r_3 f_{21}^4}{L_2 C_s}\right]^{1/3}$$

substituting in equation (8.8)

$$M_{\min} = \frac{r_s L_1}{f_1}\left[\frac{x r_3 f_{21}^4}{L_2 C_s}\right]^{1/3}.$$

The range from maximum to minimum magnification (R_g) is therefore given by:

$$R_g = \frac{L_2 L_3}{f_2 f_3 r_s}\left[\frac{x r_3 f_{21}^4}{L_2 C_s}\right]^{1/3}.$$

If a further condition is now imposed that $r_s \times$ the final stage magnification equals the maximum radius of the final image (r_s), the following substitution can be made:

$$\frac{L_3}{f_3 r_s} = \frac{1}{r_3}.$$

Hence:

$$R_g = \frac{L_2}{f_2}\left[\frac{x f_{21}{}^4}{L_2 r_3{}^2 C_s}\right]^{1/3}.$$

Now, in the weak lens approximation $C_s/f = 5[f/(S+D)]^2$. Also, f_{21}, must be approximately equal to $L_1 L_2/(L_1+L_2)$, since the second lens will be working with its image plane quite close to the objective lens in the low magnification. The above equation can, therefore, be written:

$$R_g = L_2\left[\frac{xL_1}{5r_3{}^2(L_1+L_2)f_2}\right]^{1/3}\left(\frac{f_2}{S+D}\right)^{2/3}. \qquad (8.10)$$

For maximum range, $f_2/(S+D)$ must be a minimum. It cannot be reduced below 0·3 where $V_r/(NI)^2 = 0{\cdot}0075$.

The value of f_2 is obtained through the further condition:

$$M_2 = \frac{L_2}{f_2} = \frac{M_t}{M_1 M_3}$$

where M_t is the maximum total magnification

$$f_2 = \frac{M_1 M_3 M_2}{M_t}.$$

Putting

$$M_1 = \frac{L_1}{f_1}$$

$$f_2 = \frac{L_1 M_3 L_2}{f_1 M_t}$$

we now get:

$$R_g = 1{\cdot}3\left[\frac{x f_1 M_t L_2{}^2}{r_3{}^2 M_3(L_1+L_2)}\right]^{1/3} \quad (\text{for } f/(S+D) = 0{\cdot}3).$$

The value of R_g is clearly only affected by a few percent

for values of L_1/L_2 varying from 0·5 to 2, so that within such limits we can write:

$$R_g = \left[\frac{x f_1 M_t}{r_3{}^2 M_3}.L_2\right]^{1/3}$$

$$= \left[\frac{x f_1 M_t M_3}{r_s{}^2}.L_2\right]^{1/3} \qquad (8.11)$$

for L_1 of the same order of magnitude as L_2.

From this, it is seen that the range of magnification of a three-lens instrument is not much affected by the length of the first two stages. It is in fact mainly determined by the minimum value of $f/(S+D)$ [equation (8.10)], an irreducible value. Putting typical values in the above expression, $x = 0{\cdot}1$, $f_1 = 0{\cdot}3$ cm and $M_t = 10^5$, $M_3 = 220$, $r_s = 5$ cm, and $L_2 = 15$ cm gives:

$$R_g = 74.$$

This gives a minimum magnification of 1400 with 12% distortion.

It is of interest to consider possible methods for the extension of the range without unduly lengthening the column. The simplest methods are by the use of a range of lengths of object holder, or by compensating the negative distortion of the intermediate lens against the positive distortion of the projector lens.

The use of specimen holders of different lengths does not, of course, allow a continuous range of magnification, but, for most uses, it is an adequate method. A shorter specimen holder (longer focal length) results in increased spherical and chromatic aberration; this is not of very great importance, however, since the loss of resolving power is not important at the lower magnifications obtained.

If the intermediate lens is operated so as to image a virtual image produced by the objective, then the distortion produced is negative. This negative distortion can be corrected by reducing excitation by the projector lens to

introduce an equal positive distortion. This can only be used to give an improvement of between 2 and 3 times, as higher order aberrations then arise which no longer allow the correction to be adequate. Clearly, since the diameter of image in front of the final projector is only a little under one-third of the lens bore, an absolute upper limit for the correction improvement is 3:1.

The addition of a fourth lens to the microscope gives the possibility for an increase in the magnification range. The lens might, for example, be a very weak wide bore lens immediately above or below the first projector lens. Such a lens would be used instead of the main first projector lens at the low end of the magnification range. Only a limited extension (about three times) in the range would be possible, because the pole pieces of the main first projector would limit the extent of the second intermediate image.

Another possibility is to put the extra lens coincident with the first intermediate image, and focus the rays from this into the centre of the intermediate lens. In this way, a large gain in range is possible but, in practice, a strong lens would be required and its positioning would be difficult owing to the close proximity of the first intermediate image with the objective lens.

In practice it is very seldom necessary to incorporate means for continuously varying magnification over a range wider than about 50 times and, as has been shown, this can be achieved by optimum design of the three stage system.

Practical design considerations for the projector lenses are little different from those for the objective. As has already been mentioned tolerances are less critical but higher flux densities are required. The iron circuit design must be made with especial consideration to the latter. A typical double projector lens, designed according to the principles described, is illustrated in the cross-sectional diagram of Figure 8.11. Its general construction is similar to that of the objective lens previously described and no further detailed description is necessary.

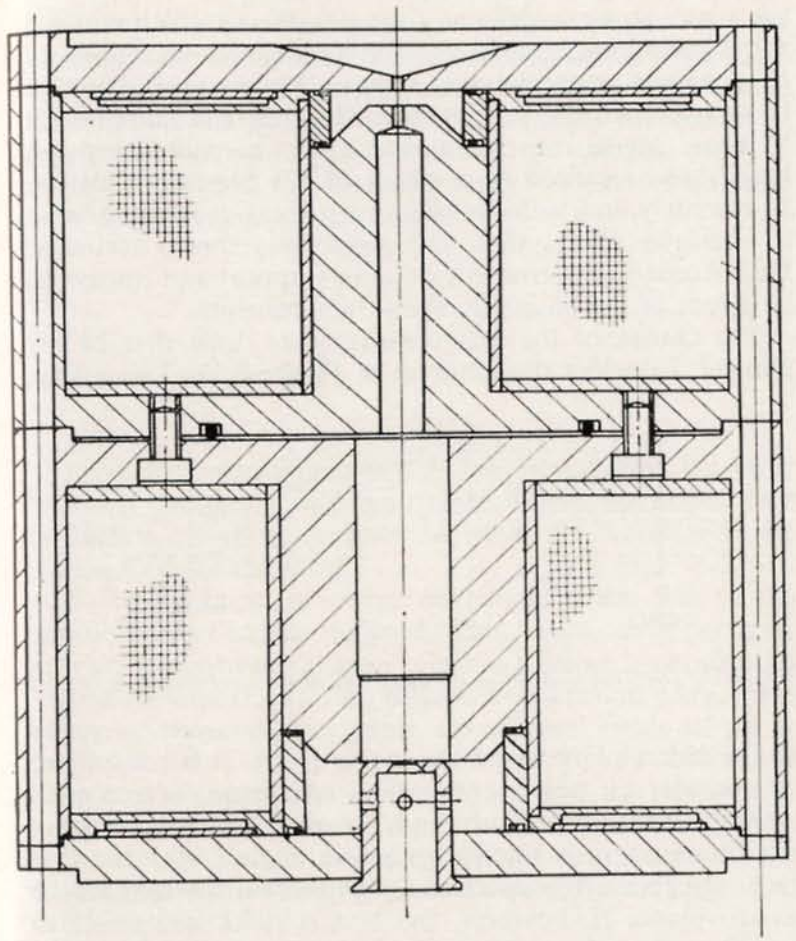


FIG. 8.11. *Cross-sectional diagram of double projector lens (M-V. E.M.6).*

Alignment of the Electron Optical System

It is not surprising that with a resolving power of atomic dimensions it is not possible to achieve adequate alignment of the electron optical system in the electron microscope without the provision of some alignment adjustments in the column. It must be stressed that, although misalignment can, in extreme cases, result in extra-axial aberrations, these

are never so important as dynamic effects which cause a misaligned image to move with very small variations of lens current or accelerating voltage. Unless an instrument is well aligned the requirement for electrical stability set by these image movements can easily be more stringent than those required as a result of the chromatic defect. Fortunately, in a well-designed instrument, it is not difficult to achieve more than adequate alignment accuracy. Unfortunately, instrument designers appear not always to be aware of the basic alignment requirements.

The causes of the image movements must first be explained. Consider the diagram of Figure 8.12: L is a lens,

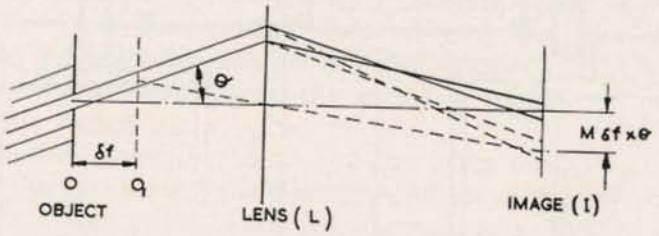


FIG. 8.12. *Diagram illustrating the alignment of the electron gun to correct 'tilt'.*

O the object plane and I the image plane. It is convenient to consider an object comprising an opaque screen with a small transparent aperture on the axis. It is assumed that the illumination is tilted through an angle θ with the lens axis. On focus, the aperture is imaged on the axis in the image plane. If, however, the lens is defocused so as to image a plane (O_1) distant δf from the object plane, the rays which are going to contribute to the image are displaced a distance $\delta f\,\theta$ from the axis in the plane O_1, and hence will be imaged at a distance $M\,\delta f\,\theta$ from the axis in the image plane where M is the magnification. Thus, a lateral shift of the whole image will take place as the focus is changed. Since perfect electrical stability can never be achieved, there will always be such lateral movement of the image in a misaligned magnetic electron microscope. In an electrostatic instrument, using unipotential lenses, the effect is absent, except that the electrons which have

lost energy in the object will be laterally displaced, thus the alignment of the electrostatic instrument is much less critical.

The alignment requirement can be defined since it is known that the stability must be sufficient to keep axial changes in focus small compared with the physical depth of focus (λ/α_0^2), and the movement must be small compared with the resolving power. Thus:

$$\delta f < \lambda/\alpha_0^2$$

but,
$$\delta f\,\theta < d.$$

Hence,
$$\theta < d\alpha_0^2/\lambda.$$

Putting $d = 0{\cdot}6\lambda/\alpha_0$ gives:

$$\underline{\theta < 0{\cdot}6\alpha_0.}$$

Thus, the requirement is that the illumination tilt must be small compared with the objective aperture angle. This is quite a stringent requirement since the aperture angle is already only about 0·3°.

A further cause of image movement arises due to the rotation effect in the magnetic lens. Even for a perfectly aligned system the rotation effect is causing a continuous loss of resolution as the voltage and current vary. This effect is, however, negligible for normal fields of view, provided the stability requirement set by the chromatic defect is maintained. If the instrument is misaligned, however, the effect can be significant, but more important is the inconvenience caused by the centre of field of view moving across the screen as, for example, the magnification is changed.

Alignment Technique

The very causes of the defects arising from misalignment provide a ready means of testing for the misalignment, and thus make correction a relatively simple matter.

In a three-stage instrument with single condenser lens, for example, the alignment conditions which must be met are as follows*:

(a) The divergent beam from the electron gun must be centered with the condenser lens aperture.

* The alignment criteria for the double condenser lens are discussed later.

(b) The focused illumination spot must be centered on the objective lens axis.
(c) The illumination must be parallel with the objective lens axis.
(d) The axis of the second projector lens must cut the screen near its centre. (This condition is easily met in manufacture and does not require the provision of a means for subsequent adjustment.)
(e) The axis of the objective lens must be aligned with the optical centre of the second projector lens so that the centre of rotation of the image, for varying objective current, is in the centre of the screen.
(f) The first projector lens must be aligned symmetrically with the objective lens axis.
(g) The objective aperture must be aligned centrally around the objective axis.

Strictly, (e) and (f) should specify that the axes of the first and second projector lens should be coincident with that of the objective but, in practice, with normal engineering tolerances on the lenses, the alignment specified is quite sufficient for even the highest performance instrument; that is to say, the objective lens axis must pass through the centres of the two projector lenses but their axes need not be exactly parallel to that of the objective lens.

It is important that sufficient alignment movements shall be provided to ensure the above. It is also important that there shall be no more than the minimum necessary movements and that they should be chosen with a view to a logical aligning sequence; otherwise a system will result which, although alignable in theory, cannot be aligned except by a tedious trial and error process. Some existing instruments still require such a trial and error process of alignment and it should be stressed that a purely logical process is quite practicable.

The method of tilting the electron beam from the gun to centre it on the condenser aperture, by means of an adjustable anode aperture plate, has been described earlier in this chapter. It is helpful to include in the instrument a

small Faraday cage which can be swung into the beam under the condenser aperture to collect the current passing through it. In this way the alignment can be achieved simply by adjusting for maximum current transmission. At the same time, since the condenser aperture defines the beam angle, and the source size is relatively constant and known, it is possible to use this measured current to calculate the brightness obtained from the electron gun and hence check on its efficient operation.

The requirement (a) is met by a simple transverse movement of the electron gun. To satisfy (b) a tilt of the electron gun is required and it is convenient to provide this tilt so that it occurs about a centre corresponding to the object plane or the source. In this way tilt may be provided without upsetting the transverse positioning. To ascertain when the tilt alignment is correct it is necessary to find the position where no transverse image movement takes place when the objective lens strength is changed. Unfortunately this test is not directly applicable because image rotation is present, as well as transverse shift. It can easily be seen that the superposition of these two image movements results in the apparent centre of rotation being displaced from the true centre. Before, therefore, correcting the illumination tilt, it is necessary to align for rotation. A method of separating the transverse movement from tilt has been suggested by the present author (1947) and is widely used. The method depends upon the fact that, if the current in the objective lens is reversed in direction no change in focus results, but the rotation angle does reverse. To apply this method an image is first obtained at low magnification (e.g. with the first projector lens off in a three-stage instrument). A distinctive marker point on the image is noted and the objective current reversed. The whole image rotates on the screen and the marker point moves to some new position. The axis of rotation lies somewhere on the perpendicular bisector of the line joining the new and old position of the marker point. This line is envisaged. The marker point is moved by the mechanical stage controls, to a position on the envisaged

line and the current reversed again. A further movement occurs unless, by chance, the marker has already been put on the true centre. A new perpendicular bisector is envisaged, as is also its intersection with the first one. This should be the correct centre. The marker point is moved to this point, by the mechanical stage, and the current again reversed. It should be found that the marker point now moves very little, and what movement remains can be reduced by a further application of the process, possibly at a higher magnification. Thus, the centre of rotation of the image is found on the final screen. It will possibly be well off centre. To bring it to the centre requires a translational movement of the top part of the instrument above the first projector lens. Applying this, the rotation centre is brought to the screen centre. The objective axis now passes through the centre of the second projector lens.

If, now, a change in objective lens excitation current causes a movement of the centre of the image, it must be due to illumination tilt. To correct this the objective lens is defocused as far as possible and the marker point, now off centre, is brought back to the centre by using the electron gun tilt control. It will then be found that no movement occurs on taking the objective lens through focus. The illuminating beam is now aligned with the objective axis.

The alignment of the first projector lens can be achieved by imparting to it a translational movement, or by including means for tilting the upper half of the column with respect to the projector lenses. In order to prevent this tilt from upsetting the objective second projector alignment, it must be centered on the centre of the second projector lens so that the tilt does not move the point of intersection of the objective axis with the plane in which the second projector lens lies, this point being on the centre of the second projector lens.

The alignment process is simple. When the first projector lens is switched on the marker point will move off centre if the lens is misaligned. It is brought back to the centre by the above-mentioned tilt control; after this it should remain centered at all magnifications whether the first projector lens is on or off.

It is usually desirable to go through the whole alignment process twice to satisfy the more exacting requirements. Once the main lenses are aligned they should not require further adjustment unless the instrument column is dismantled. The gun tilt and shift controls will usually require day to day adjustment.

It sometimes happens that an instrument can be aligned to give no image movement when the objective current is changed, but will give a movement when the high voltage changes by small amounts. This results from the presence of a stray magnetic field component, perpendicular to the axis, usually in the region of the objective lens, or immediately following it. In the presence of such an effect perfect alignment cannot be achieved and resolving power may well be limited by image movement due to voltage or current fluctuations, and the resulting misalignment. Experience has shown that care in the design of the magnetic system of the lens can eliminate any such defect, though it may be worthy to note that the present author only achieved this very desirable result by eliminating lenses with separate pole pieces as has been discussed earlier in this chapter.

The above alignment process is best carried out with the objective aperture removed. To align the objective aperture it is inserted into an approximate position and the intermediate lens adjusted to the diffraction condition so that it images the back focal plane of the objective lens, and hence the objective aperture, on to the screen. This should be carried out with a specimen, such as a thin support film, giving some electron scatter, and with the condenser lens defocused to give a near parallel illuminating beam. The image will appear as a small circular illuminated spot with an intense point of light where the parallel illuminating beam converges to a focus. In general this bright point may be off-centre, or even outside the field of view. Centering is achieved by moving the aperture until the bright point of light is central within the aperture image. This operation is extremely quick and simple, especially if a 10 times magnifier is used to observe the image. If no scattering occurs in the object plane the bright spot of light will

appear, but the aperture itself will not be illuminated and it becomes difficult to see the position of the aperture around the spot.

Alignment Requirements for Double Condenser Lens

It is clearly undesirable to have to introduce additional aligning mechanisms with the double condenser lens system. Possible misalignments between the source and the two components of the lens can be split into three components as follows:

(a) The source may be off the axis of the first lens.
(b) The axis of the first lens may be off the centre of the second lens and
(c) The axis of the second lens may be off the centre of the first lens.

These three defects can result in two types of alignment error. In the first place the illumination may be lost when the spot is defocused by changing the excitation of the second lens, secondly the illumination may be lost when the magnification is changed by variation of the excitation of the first lens. The first defect is of more importance as it occurs during a frequently occurring operation. The second defect is likely to be annoying but need not be considered serious as long as the spot movement, occurring over the full change in de-magnification, does not exceed about 200 microns. A movement such as this can readily be corrected by the transverse shift control to re-align the illumination.

The first defect mentioned in the previous paragraph, the loss of illumination on defocusing, can also occur with a single condenser lens. However, the alignment requirement necessary to avoid its occurrence is not stringent. The requirement is that the off-axis distance of the source should not exceed about 1 mm. The second condition is much more stringent and analysis* shows that, for no loss of illumination, the source should never be off axis by a distance greater than its own radius. However, as has been mentioned before, since the illumination magnification is

* Not yet published.

not varied very frequently, and the error, if not too big, can be simply corrected by the gun transverse shift adjustment, a tolerance of about 10 times the source size can be tolerated, i.e. approximately 200 microns.

The alignment tolerances (b) and (c), described earlier, are of the same order as those necessary in the cases of the objective and projector lenses and can be met by established manufacturing techniques.

Thus, provided the comparatively minor inconvenience of re-aligning, when large changes in source magnification are made, can be tolerated, and provided the same order of accuracy in manufacture can be maintained in the double condenser lens as in other lenses, there is no real necessity for additional alignment adjustments in the case of the two-stage condenser lens.

Alignment Mechanisms

It has been explained in the previous section that the main alignment adjustments necessary in a three-stage instrument are:

1. An adjustment to centre the beam from the electron gun about the condenser lens aperture.
2. Lateral movement of the illuminating system with respect to the objective lens.
3. Tilt of the illuminating system about a centre at the object.
4. Lateral movement of the objective lens, with the illuminating system, with respect to the first projector lens.
5. Tilt of the objective lens, with the illuminating system, about a centre in the second projector lens.
6. Centering adjustment for the objective aperture.

The illuminating system may be aligned mechanically or magnetically. The mechanical systems usually comprise a pair of sliding surfaces with appropriate adjusting screws, one pair of surfaces being flat and the other spherical, with a centre at the object. Vacuum sealing between the two parts of the tube is usually achieved by flexible metal bellows. A

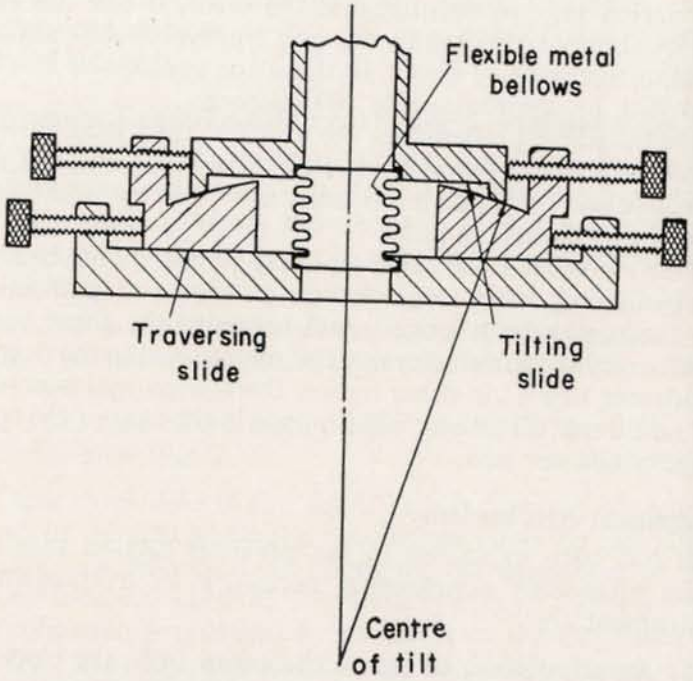


FIG. 8.13. *Mechanical alignment section for electron gun (M-V. E.M.3).*

typical arrangement is shown in the cross-sectional diagram of Figure 8.13. A large angle of tilt is required (20°) in the case of instruments designed for reflection working. To achieve this with a modern double condenser lens system, which necessitates a pumping manifold connection direct to the electron gun, the mechanical system becomes very cumbersome, especially since the small illuminating spot requires very high mechanical stability. A system making use of combined magnetic and electrostatic deflection is incorporated in the M-V. E.M.6 instrument (Haine, Agar and Mulvey, 1958). This is illustrated diagrammatically in Figure 8.14. It comprises two electromagnets with their associated iron circuits providing deflecting fields. The pole faces of these magnets are electrically insulated and are used also as electrostatic deflector plates, to produce a

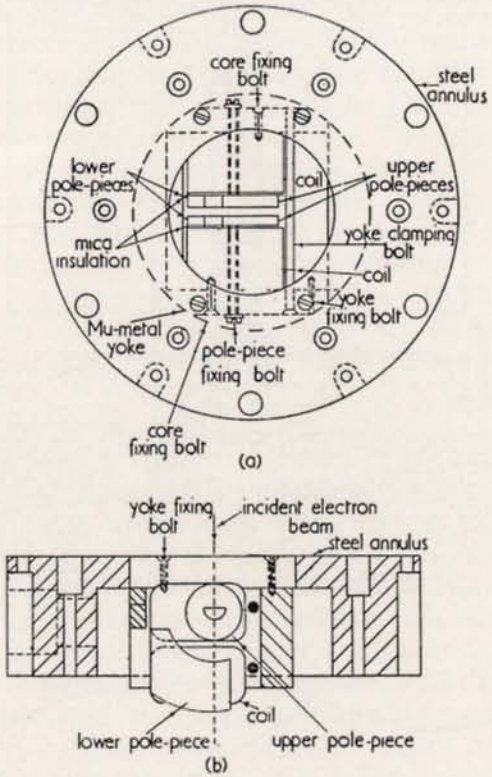


FIG. 8.14. *Diagram illustrating the magnetic cum electrostatic illumination alignment section.*

deflection perpendicular to that resulting from the magnetic field. It will be seen that a shift of the illuminating spot can be obtained by the excitation of one magnet or electrostatic deflector. To produce a tilt the first deflector bends the beam away from the axis and the second, deflecting in opposite sense, deflects it back towards the axis. For normal alignment, in transmission working, such an arrangement presents no design difficulties. When, however, such a system is used to tilt the beam through a large angle, careful design is necessary to avoid the introduction of a large amount of astigmatism. In addition, since it is highly

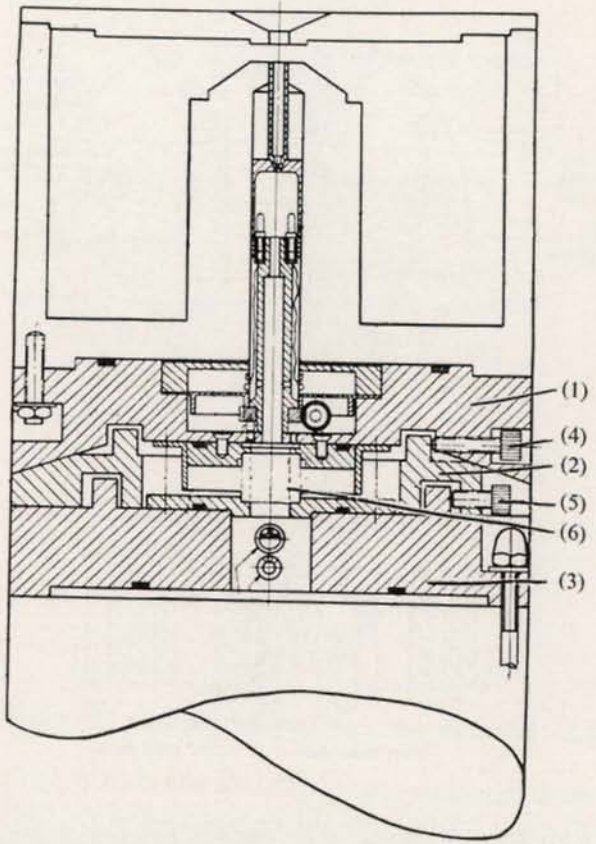


FIG. 8.15. *Cross-sectional diagram of the intermediate alignment section (M-V. E.M.6).*

desirable that the tilt shall be adjusted by one control, it is necessary that the deflections shall 'track' so that no movement of the illuminating spot accompanies the change in tilt. Theoretical aspects of this system have been discussed by Archard and Mulvey (1958) and have led to the design of pole pieces described in the paper previously referred to.

The intermediate aligning mechanism is usually mechanical and must maintain the very high stability necessary in the instrument columns. A typical alignment stage is shown in Figure 8.15. It consists of the plate (1) on to which

is bolted the objective lens. The underside of this plate has a concave spherical surface which slides on a matching convex spherical surface on the plate (2) which, in turn, slides on the upper flat surface of the plate (3), which is bolted to the projector lens assembly. The adjustment for tilt is achieved by the four adjusting screws (4) while that for the lateral movement is achieved by the adjusting screws (5). The sliding surfaces are lapped together to achieve a good fit, and the whole assembly is fitted with four clamping screws (not shown in the figure) which can be tightened down after alignment to ensure rigidity. Vacuum sealing is by the flexible metal bellows (6).

Objective Aperture

The objective aperture in the electron microscope is included mainly for the enhancement of contrast. The diameter of the apertures used varies from 20 to 50 microns depending upon the focal length of the objective lens and other operational conditions. The aperture is located in the back focal plane of the objective lens and must be accurately centered on the axis of the instrument. A most important requirement for the aperture is that it should be perfectly clean and free from any insulating contamination. This has, in the past, proved a major source of trouble in high resolving power work. For many years platinum aperture discs were used, since these could be cleaned by strong acid or caustic treatment. However, as has been mentioned elsewhere, it was found that platinum and platinum alloys were apt to have an extremely resistant surface film which caused astigmatism in the imaging beam. Even the most rigorous cleaning treatment did not always remove this film. Haine and Mulvey (1954) showed that a molybdenum aperture disc could be cleaned easily by flashing it at about 1200°C, *in vacuo*, and such discs are now in common use and are available commercially.

It is desirable that the aperture disc should be mounted in a mechanical stage, to allow centering, and should be readily removable from the instrument for cleaning or replacement. Various stage designs have been used for this

purpose, one of these is the simple rod type system illustrated in the photograph of Figure 8.16 which shows the objective lens with the top plate removed. The mechanism comprises a rod, containing recesses for three interchangeable apertures, which slides into the space between the upper and lower pole faces through a double rubber gland seal. The end of the rod screws into an extension which, in the operating position, is held in a V groove under

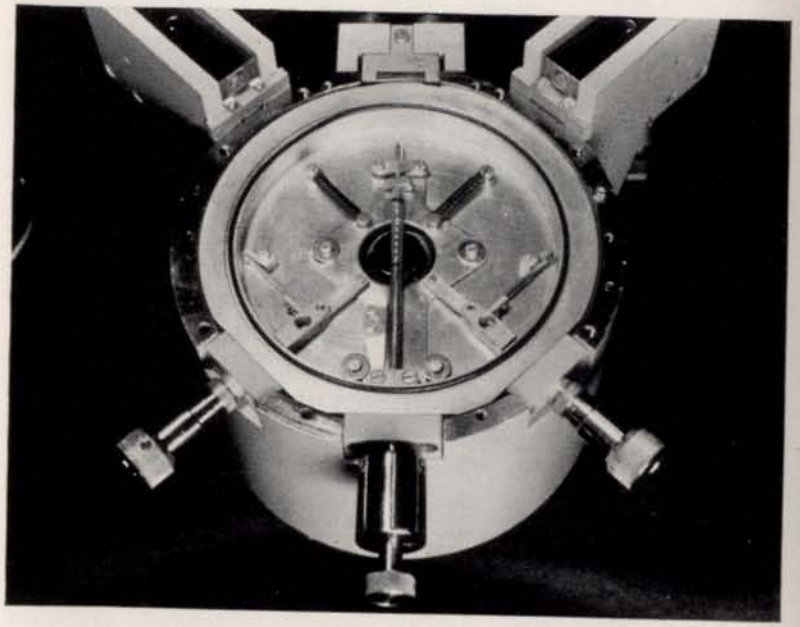


FIG. 8.16. *Photograph of the objective aperture alignment mechanism with the top plate of the objective lens removed* (*M-V. E.M.6*).

a leaf spring. A similar V groove locates the other end of the rod. These V grooves are mounted on a stage plate which can be moved, for centering, by the two differential drive screws shown. To withdraw the aperture, the rod, with its extension, is withdrawn through a double gland seal until the 'O' ring, located in a groove in the extension rod, enters the withdrawal tube. The rod can then be

unscrewed from the extension and removed from the instrument without disturbing the vacuum.

Magnetic Screening

The presence of a stray transverse magnetic field will deflect the imaging beam in the electron microscope. The effect of a steady, non-time varying field is to deflect the beam by an amount which will vary with the accelerating potential on the electron gun, but will not vary with the objective lens excitation current. For a given value of accelerating potential the affect can be corrected by the alignment procedure but, as has been explained in the section on alignment, it results in a non-coincidence between the illuminating beam tilt correction for variations in high voltage and for variations in the objective current. Such a field can thus prevent the attainment of the necessary alignment accuracy. In practice it is rarely found that such static fields arise from sources external to the instrument. If present they will inevitably arise from asymmetric saturation effects within the instrument. The main external source is the earth's field and this will normally be adequately shielded by the iron parts of the instrument.

Time varying fields have more serious effects and are a common source of trouble. Only the component of field perpendicular to the electron beam is important, and the most sensitive part of the electron optical system is that immediately following the objective lens. In this region fields of the order of 10^{-7} gauss can deflect the beam by an amount sufficient to limit the resolution to a few Å. Since, in a normal laboratory building, stray magnetic fields of power frequency will often exist with an amplitude between 0·1 and 1 milligauss, it is necessary to achieve a magnetic screening on the microscope which will reduce fields in the ratio of 10^4–10^5:1. Many instruments incorporate mu-metal screens for this purpose, but in others it has been found that the use of soft iron or low carbon steel in the fabrication of the various parts of the microscope column will provide an adequate screening factor.

The Object Chamber and Mechanical Stage

The evolution of the modern, highly stable mechanical stage has too complex a history to describe in detail. The basic problem is too difficult to solve by logical kinematic engineering methods. Indeed, the application of such methods tends almost to prove the impossibility of attaining a solution. Different designers have achieved a high degree of success, if not perfection, in different ways.

It is required to move the specimen over a distance of about a quarter of a centimetre in two co-ordinate directions, to set its position to about one-tenth of a micron and, once having set it, for it to remain stationary within about 1 Å for a period of up to 30 seconds. The movements from the two-stage controls should be approximately orthogonal and free from backlash greater than about 0·1 micron. A satisfactory design has only been achieved because the electron microscope itself allows observation of the effects of experimental rearrangements with the required high accuracy. Successful stages are all simple in principle. They are far from ideal kinematic systems and rely very largely on a careful balance between freedom of movement and frictional restraint. Friction is essential to damp vibrations, rigidity is essential to prevent backlash and undesirable drift. In addition, some provision for maximum heat dissipation from the specimen and objective lens must be provided, to reduce temperature changes which result in slow drift of the stage, due to thermal expansion.

There has been very little detailed description of the design considerations for a stage in the literature, and very few really satisfactory designs have been achieved. It may be of value to describe some of the conclusions which have been reached in the author's experience as relevant to the design of a satisfactory stage for really high resolution working.

The specimen support grid is usually a 200-mesh circular disc with square apertures. It is important that the metal used should have high thermal conductivity and copper is normally employed. Available grids are manufactured by a

photo-mechanical process. The grids must be mounted on some form of specimen holder with good thermal contact. It has been usual to mount them by a cap which slides on to the tubular end of the specimen capsule, being held by friction. This arrangement often gives poor thermal contact and a screwed cap is preferable, though it leads to a larger total diameter requiring a wider bore in the lens pole piece.

The specimen capsule has to be inserted into the mechanical stage, usually through an air lock, and must be held in place rigidly and with good thermal contact. This is not easily achieved. The necessity of using an air lock makes it difficult to provide a very positive clamping arrangement to hold the capsule in place, especially since the capsule must be entirely disengaged from the air lock mechanism to allow freedom of movement with the mechanical stage. The author has used a conical seating between capsule and stage. This has proved satisfactory provided the cones are carefully lapped together to provide perfect fit. Even with moderate lapping sufficient for good mechanical rigidity the heat conduction can be very severely limited. This difficulty is greatly increased because, in the vacuum, no heat transfer other than by conduction or radiation can occur. In one experiment a conical fit between a copper capsule and copper stage, with good mechanical rigidity, led to a temperature drop of some 60°C across the joint. This was reduced to a very small value by improved lapping. A disadvantage of this arrangement is that interchangeable capsules are hardly possible.

The stage itself must be rigid and of good thermal conduction. It should slide, preferably, directly on the top of the lens. It should 'sit stable' on the sliding surface, which means it should contact this surface at three points as far from the centre as possible. Since point contacts lead to rapid wear and poor thermal conductivity it is desirable to provide three pads of reasonable area. The pads should be lapped on to the lens to provide accurate registration and free movement. All sliding surfaces *in vacuo* are subject to much higher friction than in air and the provision of lubrication is difficult, as it is highly

undesirable to have any organic material in the microscope. Some advantage can be gained by the use of such solid lubricants as molybdenum disulphide, but their effect is not lasting. Hard steel pads sliding on the soft iron of the lens prove fairly satisfactory.

The stage must be held in contact with the lens by some form of spring loading which does not itself produce excessive restraint to free movement. Two leaf springs with

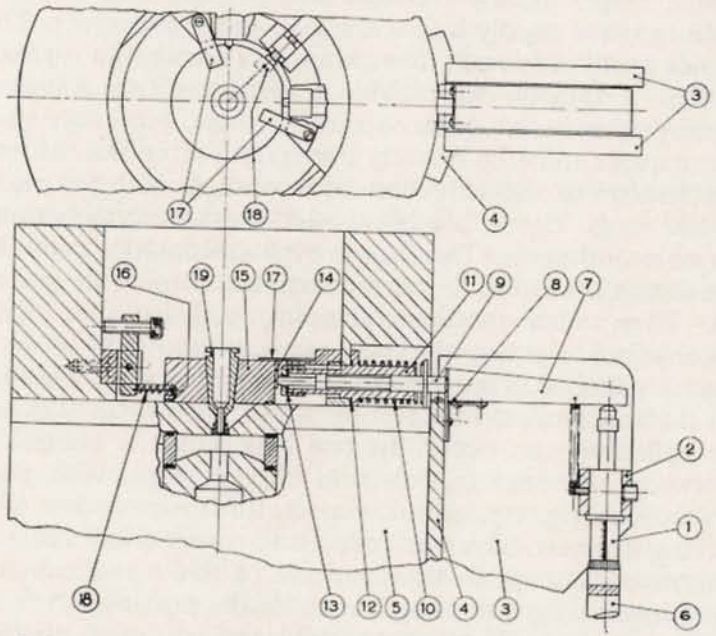


FIG. 8.17. *Cross-sectional diagram of an experimental mechanical stage system.*

carefully adjusted loading have proved satisfactory here. Beryllium copper is a useful non-magnetic material for such springs. Figure 8.17 shows a cross-sectional diagram of the arrangement used by the author in the experimental prototype of the M-V. E.M.6.

The driving lead screw comprises a standard micrometer screw (1). This is clamped rigidly into a block (2), mounted between two plates (3), and welded to a third plate (4). This

is fitted and bolted directly on the side of the objective lens (5) (see also the photograph of Figure 8.16). By mounting the main bracket on to the objective lens, rather than on to the specimen chamber, any movement of the object chamber relative to the lens does not affect the position of the object relative to the lens. The micrometer is driven from a control knob conveniently placed at the side of the instrument just above the desk level via the drive shaft (6). The movement of the micrometer is transmitted to the mechanical stage via the lever (7) with crossed spring support bearings to eliminate friction and backlash. Rotation of this lever takes place about the point (8) and the end of the lever bears on to a push rod (9) giving a lever ratio of about 10:1. The push rod is held in contact with the lever by means of the compression spring (10) and slides into the tube (11) with bearings at either end, the diameter of the rod being relieved at the centre to reduce friction.

The tube (11) is vacuum sealed into a recess in the object chamber by the rubber gasket (12). The end of the push rod enters the vacuum through the rubber gland seal (13) and bears on to a hard steel plate (14) rigidly attached to the side of the stage (15). The stage slides on the lapped upper surface of the objective lens on a hard steel ring (16) which is inserted into the bottom of the copper mechanical stage block; the surface of the ring is relieved, except for three short arcs, spaced 120° apart, to form pads. The leaf springs (17) are carefully adjusted to give the right pressure to hold the stage in contact with the object lens.

Two such drives displaced at 90° are, of course, incorporated. The stage is held against the push rods by means of the adjustable tension spring (18), which is located at as low a level on the stage as possible in order to minimise any torque which would tend to lift one side of the stage.

The specimen capsule (19) is lapped into a conical hole in the copper block. The specimen grid is held on the end of the capsule by a screwed cap.

A modified version of this design is used on the M-V.

E.M.6 in which provision is made for stereoscopy by tilting the centre part of the mechanical stage, holding the specimen capsule, against cylindrical surfaces whose axis passes through the normal position of the specimen. In this way tilt can be provided without lateral movement, facilitating the taking of stereoscopic pairs.

The above design of mechanical stage has proved in practice to have a rigidity and freedom from backlash at least an order of magnitude better than previous designs. The main contributory reasons for this are the extremely rigid brackets supporting the drive screw, the high quality screw, the improved fulcrum arrangement for the lever, the careful balancing of the spring loading and friction of the stage, the incorporation of the hard steel sliding ring, and the mounting of the bracket on the lens rather than on the object chamber. Thus it is in the attention to detail rather than in the basic design that improvement may be obtained.

It will be appreciated that one difficulty in experimenting with a mechanical system of this sort is to devise means for ascertaining which part of the system is responsible for any defect noticed in its operation. It is fairly easy to ascertain whether defects in the mounting of the bracket or lever are responsible for incorrect operation by applying small pressures to these components and observing the sensitivity as judged by the resulting movement of the microscope image. In developing the stage system described above it was found that this method could be extended to the parts of the stage inside the vacuum by using a small manipulating rod, inserted into the object chamber via a flexible metal bellows, to apply small forces to different parts of the stage.

The different types of airlock for insertion of the specimen into the vacuum have been adequately described in the literature and need not be commented on in detail here. The inclusion of an airlock in the object chamber is of importance both to allow quick changing of specimens and also to avoid the repeated adsorption of water vapour on to the surfaces inside the instrument, thus leading to short filament lives through water vapour attack. Apart from reliability and ease of operation, the most important design

consideration for the airlock is that it should not detract from the satisfactory operation of the mechanical stage. It is convenient to mount the airlock mechanism on one or more flanges on the side of the object chamber so that it can be easily removed, to allow ready access for adjustment of the mechanical stage, or for the insertion of any special attachments such as a reflexion stage. This arrangement is typified in the M-V. type E.M.6 instrument where not

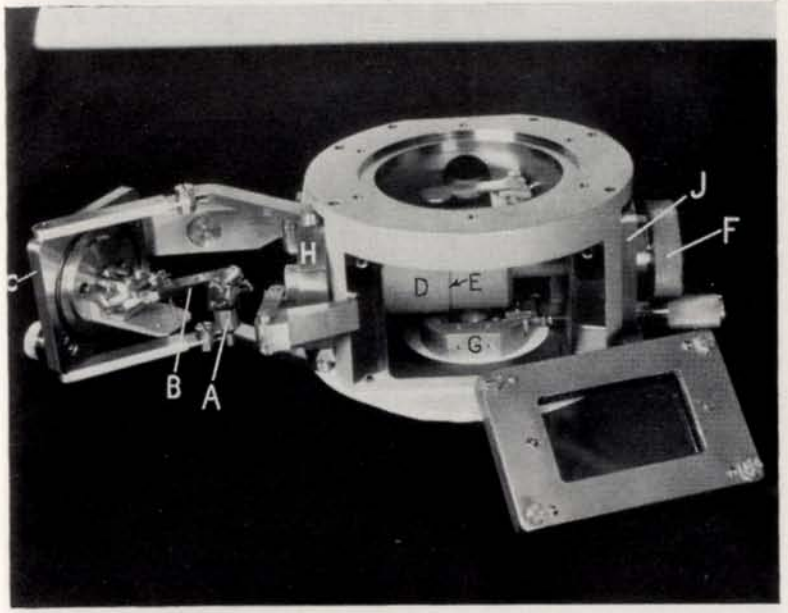


FIG. 8.18. *Photograph illustrating the specimen airlock system (M-V. E.M.6).*

only can the normal transmission stage and the airlock system be quickly removed and replaced by a reflexion stage but a large window is provided allowing easy observation of the operation of the airlock system. Figure 8.18 shows a photograph of the specimen chamber of this instrument. The specimen capsule (A) is inserted into a double hook mounted on the end of the lever (B) which can be raised and lowered by an external control knob

mounted on the flange (C). This assembly is inserted into the side of the chamber so that the specimen capsule hangs inside the tubular member (D). This tubular member is in two parts, vacuum sealed together at (E). After insertion the tube is rough pumped and the right-hand half is withdrawn, by means of a lead screw operated by the control knob (F). The airlock chamber is now open to the main column and the specimen can be lowered, by tilting the supporting lever, into the mechanical stage (G). The lever control is so arranged that its positive downward pressure can be applied to the specimen to ensure that it seats firmly in its conical seating. This pressure is, of course, applied and then released before subsequent use of the mechanical stage. To remove the specimen the lever is lifted, withdrawing the specimen capsule from the stage, the airlock tube is closed and, after letting air into the airlock tube the flange is hinged away from the instrument.

The figure shows the front flange of the specimen chamber, with its viewing window, removed. To replace the transmission stage by that for reflexion the two side flanges (H) and (J) and the front plate are removed, the side flanges are replaced by blank plates and the reflexion stage is inserted instead of the front plate. The reflection stage is shown in the photograph of Figure 8.19. The specimen is mounted on the screw (a) and provision is made for tilting it on a horizontal axis passing through its face, for rotating it, for moving it transversely, and for a transverse movement perpendicular to the specimen face. The rotation and tilt controls are calibrated.

Mechanical Design of Column

The following are comments on some of the more important general design details of the microscope column.

It has already been stressed that the rigidity of the microscope column is essential to good performance. Experience has shown that adequate rigidity can be obtained if attention is paid to the important design details, and if ambient vibration is not excessive. A more critical part of the microscope column is that immediately following the

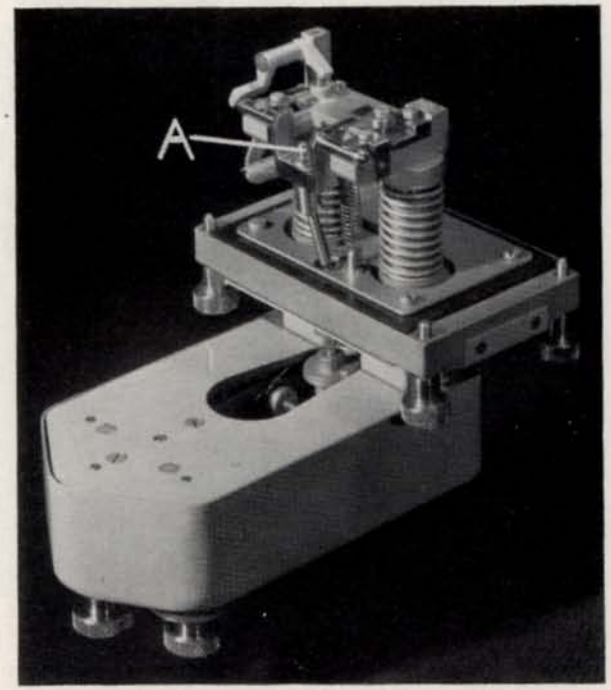


FIG. 8.19. *Photograph of reflection stage system* (*M-V. E.M.6*).

objective lens, and joints between sections and within sections are the main cause of column flexing. To minimise the effects the column joints should all be metal to metal, that is, the sealing gaskets should be compressed until metal-metal contact is assured. This contact should furthermore be concentrated near the outer diameter of the sections. It is desirable to recess sections so that contact is assured on the outer periphery.

It will be clear that the intermediate alignment section is the most critical assembly, since the joints here must be capable of sliding one over the other while, at the same time, retaining rigidity. A typical design for this section has already been described. To ensure rigidity the meeting surfaces are arranged around the periphery and are lapped

together. It is of interest to note that the stability of the column can be fairly well assessed simply by pressing on it with the hands, or by rocking the instrument. This is, of course, a somewhat empirical test but nonetheless effective. A faulty joint is easily detected by squeezing the joints in turn to find the relative sensitivity, i.e. the relative amount of image movement produced at each joint. In a well-designed and constructed instrument the maximum bending torque which can be applied to a column by twisting it with the hands should not produce an image movement of more than about 1 cm when operating at a magnification of 100,000 ×.

The Camera

The camera of the electron microscope is mainly a mechanism for inserting photographic plates into the vacuum, usually via an airlock, and for moving them into position under the fluorescent screen. Many designs of camera have been described with single- or multiple-plate facilities and with or without airlocks. There has been considerable controversy concerning the best size and shape of plate to be used and even whether to use plate or film. It would serve no useful purpose to go over the arguments here since they are very much a matter of individual preference, and requirements may depend on the type of work to be undertaken. Descriptions of the mechanical arrangement and the methods of operation are given in the various makers' catalogues as well as in the published literature.

A typical camera assembly is shown in the photograph Figure 8.20. This camera holds six photographic plates, size $4\frac{1}{4}$ in. × $3\frac{1}{4}$ in. Each photographic plate is contained in a separate light-tight cassette. The cassettes are arranged in a stack and, on operation of the control lever on the front of the camera, the cassettes are pushed, one by one, into position under the fluorescent screen. Their lids are retained during this motion and reclosed as the cassette is withdrawn. The assembly fits into a vacuum box under the microscope which can be isolated from the main column by an airlock

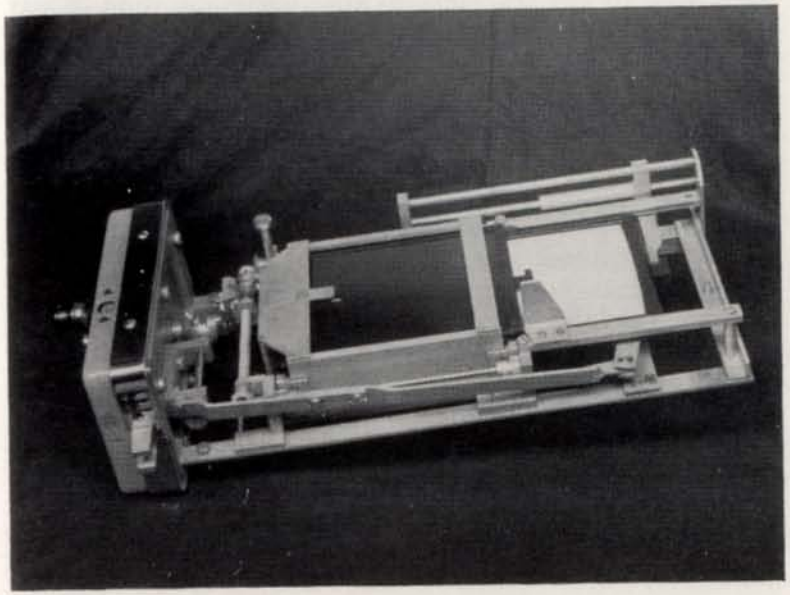


FIG. 8.20. *Photograph of electron microscope camera* (*M-V. E.M.6*).

plate. As for the specimen airlock, the vacuum chamber can be pumped to a rough vacuum by the mechanical vacuum pump before opening the sealing plate. With such a camera, between one and six exposures can thus be made without the necessity for changing the plates. An attractive feature of this design is that the working parts, all being attached to the removable front plate, are easily accessible for maintenance.

Vacuum Joints

The many demountable vacuum joints in the electron microscope are, nowadays, mostly sealed with moulded rubber 'O' rings. These rings are inserted in grooves and are usually arranged so that, when the flanges are clamped together, they come into metal-metal contact by the time just sufficient compression of the gasket is obtained to ensure a good seal (10–15 thousandths of an inch). If the seal is correctly designed and made, it is not necessary to

use grease on the 'O' rings to ensure vacuum tight joints; indeed, it is undesirable to use grease anywhere in the electron microscope as it leads to excessive specimen contamination and instability due to static charges.

A rubber gasket joint can normally be demounted many tens of times without the necessity of replacing the rubber gasket. It is important that the surface of the gasket should be kept perfectly clean and free from any lint or such matter which may provide a gap for air leaks. Similarly the surface of the metal on to which the gasket will seal must be clean and smooth and great care must be taken to avoid it becoming scratched. By careful attention to such details vacuum leaks can become non-existent.

In certain cases rectangular vacuum seals are necessary and here the gaskets are normally cut from sheet, or may be specially moulded. The requirements for avoiding leaks are no different from those described above for circular seals though special care must be taken to see that the flanges are clamped down uniformly all round.

Rotating seals into the vacuum also make use of 'O' rings or gland washers cut from sheet. If the shaft is highly polished and the rubber in contact with the shaft is smooth, a good seal can be maintained without introducing much more friction than is normally present in passing the shaft through an unlubricated bearing. It is usual to make the hole through the gasket about 10 to 15 thousandths of an inch smaller than the shaft so as to produce the necessary pressure for sealing. It is desirable that the rubber used in gland seals shall be made hard by heavily loading with carbon. The use of a soft rubber leads to excessive friction and wear.

REFERENCES

ARCHARD, G. D. and MULVEY, T. *J. Sci. Instrum.*, **35** (1958), 279.
HAINE, M. E. *J. Sci. Instrum.*, **24** (1947), 61.
HAINE, M. E. *J.I.E.E.* (*London*), **94** (1947), 447.
HAINE, M. E. *Proc. Internat. Conf. on Electron Microscopy*, London (1954). Published by the Royal Microscopical Society.
HAINE, M. E., AGAR, A. W. and MULVEY, T. *J. Sci. Instrum.*, **35** (1958), 357.

HAINE, M. E. and MULVEY, T. *Proc. Internat. Conf. on Electron Microscopy*, London (1954). Published by the Royal Microscopical Society.
HAINE, M. E. and PAGE, R. S. *Electron Microscopy Proc. Stockholm Conference*, Almqvst and Wiksell, Stockholm (1956).
LIEBMANN, G. *Proc. Phys. Soc.* (London), **65** (1952), 94.
LIEBMANN, G. *Proc. Phys. Soc.* (London), **66B** (1953), 448.
MULVEY, T. *Proc. Phys. Soc.* (London), **66B** (1953), 441.
LE POOLE. Unpublished.
RUSKA, E. *Proc. Internat. Conf. on Electron Microscopy*, London (1954). Published by the Royal Microscopical Society.

CHAPTER IX

RELATED TECHNIQUES AND INSTRUMENTS

DURING the course of the development of the electron microscope a number of related techniques have been evolved which should be mentioned, at least briefly, in any exposition on the main instrument. It is not intended to analyse these techniques in any detail. Their application is, in each case, very much more restricted than the main transmission instrument. Nevertheless, they have their important place in a rapidly expanding field of instrumentation. All the modified techniques are aimed at providing additional information on microstructure. There are three general types of information which can be obtained by electron (or ion) interaction, microscopic, atomic structure and elemental analysis. Microscopic information may be obtained by imaging or by probe beam scanning; atomic structure may be examined by diffraction, and by the study of characteristic X-ray emission, the elements may be identified and their relative abundance estimated.

The first two techniques to be mentioned are applied directly to the transmission instrument. The rest involve special instruments which, though in some cases very similar to the transmission instrument, are sufficiently different in basic design to render it difficult, uneconomic, or inefficient to incorporate the technique in commercial transmission instruments. It would, indeed, be possible to devise an instrument capable of dealing with most of the techniques and much thought has been given to this possibility. The conclusion which is inevitably reached is that a very expensive compromise must result, in which

much would be sacrificed in the performance of the basic instrument.

The Electron Reflexion Technique

If an electron beam is allowed to fall at near glancing incidence on a solid surface, a reasonable proportion of the electrons are scattered clear of the surface and may be used to form an image, as illustrated in Figure 9.1. The advantages of the method lie in its ability to be used

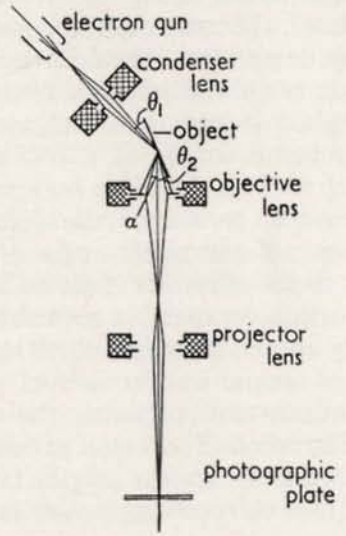


FIG. 9.1. *Schematic representation of reflexion electron microscope system.*

directly on bulk specimens without the use of intermediate replicas, and in the high contrast obtained where the slopes of the surface asperities or hollows are small. On the other hand, the method is limited in resolution because of the wide spread in energy losses of the scattered electrons.

The method was first attempted by Ruska (1933) who used an angle of 90° between the incident and scattered beams. The resolution obtained was very poor but later von Borries and Janzen (1941) reduced the deviation angle to a few degrees and obtained a resolution of a few hundred

Å. In the last few years a good deal of work has been carried out on this technique in Russia (Kushnir, etc., 1951) and France (Fert and Raporte, 1952, Fert, 1954, Fert, etc., 1955, Fert, 1955) as well as Britain (Haine and Hirst, 1953; Menter, 1952, 1953).

The results obtained by the reflexion technique depend, to a large extent, upon the two angles θ_1 and θ_2 (Figure 9.1), these being the angle between the illuminating beam and the specimen surface, and the angle between the specimen surface and the instrument axis. As would be expected from scattering theory, the intensity increases as the scatter angle $(\theta_1 + \theta_2)$ is made smaller. The contrast depends very largely on the value of θ_1 and in some cases, for example that of a well-polished surface, the contrast may increase markedly between values of θ_1 of a few degrees and a fraction of one degree. It is desirable to keep the value of θ_2 as large as possible, to reduce the picture distortion which arises, because of the small angle of viewing. The tendency has been to increase this angle to as much as 30° to reduce this distortion, or even, in some cases, to correct distortion optically in the photograph. This is, however, not so important as would appear at first sight provided that, when interpreting the pictures, the conditions of imaging are borne in mind. The value of energy spread of the scattered electrons is in the region of 100–200 V (Kushnir, etc., 1951). The resolving power is given by:

$$d = C_c \alpha \, \Delta V / V.$$

Where C_c is the chromatic constant of the objective lens, α is the objective aperture semi-angle and $\Delta V/V$ is the fractional energy spread. This leads to a resolving power in the region of a few hundred Å, though it should be possible to reduce this to about 20 Å if a lens of minimum chromatic aberration and very small aperture were used (Haine and Hirst, 1953). This, however, leads to the use of an extremely short focal length lens and very small beam limiting aperture (about 2·5 microns). A further disadvantage would be that the specimen would now lie in the magnetic field of the lens and image intensities would be very low. The

intensity of the scattered electrons is at least of an order of magnitude less than that obtained in the transmission instrument, thus leading to very low intensity images anyway. The use of a large viewing angle (θ_2) reduces intensity still further.

Because of the oblique angle of the specimen only a restricted area lies within the depth of focus of the instrument. This leads to a restricted field of view within which the image is sharp. For typical values of resolution obtained, the field of view in the direction of tilt amounts to about 100 resolved object points only. The mechanism of contrast formation in the reflexion technique depends mainly on the shadowing effects of the illuminating electron beam. They are discussed by Haine and Hirst (1953).

The application of the reflexion technique in the microscope requires, in addition to the normal alignment adjustments of the electron gun, means for tilting the gun to the angle $(\theta_1 + \theta_2)$ together with a special stage to hold the specimen. This stage should preferably have the facilities for an adjustable tilt of the specimen (to set θ_2), an adjustment to move the specimen towards or away from the electron beam to determine the axial position where the beam strikes the specimen, a perpendicular movement to search the specimen laterally, and a rotation of the specimen about the point of observation. Typical examples whereby these requirements may be obtained are described in Chapter VIII.

Limited Area Electron Diffraction

An electron diffraction pattern constitutes a Fourier transform of an object illuminated with coherent radiation. The intensity distribution in the back focal plane of the microscope objective lens gives just such a Fourier transform and therefore contains the diffraction pattern when the specimen is coherently illuminated. It is possible to project this diffraction pattern on to the final screen, or the photographic plate in the electron microscope, by an intermediate lens focusing it into the image plane of the final projection

lens which, in turn, projects a magnified image on to the fluorescent screen. In a three-stage electron microscope the intermediate projector lens may be used in this way by sufficiently reducing its excitation current (Haine, Page and Garfitt, 1950). To perform this function the objective aperture must be removed, since this would otherwise cut off the diffraction pattern. If a field aperture is inserted before the intermediate projector lens so that, under normal microscope conditions, it limits the field of view in the final image, then the diffraction pattern obtained will be only from that part of the specimen which is included in this limited field of view. Since it is possible to insert field limiting apertures as small as 25 microns, and the field of view defined by it has an extent given by the aperture diameter divided by the objective lens magnification, it is possible to obtain diffraction patterns from areas well under a micron in diameter.

Using this technique it is possible to take a specimen consisting of mixed crystals and to identify them individually by electron diffraction techniques. Most modern instruments are fitted with facilities for carrying out this function. The method can be applied in the reflexion technique also. An interesting result, which throws much light on the reflexion diffraction mechanism, is obtained if the objective aperture is adjusted so as to transmit only, or mainly, one of the diffracted beams. On adjusting the microscope to give an image, only the crystalline regions contributing to the diffracted beam are illuminated (Haine and Hirst, 1953).

Electron Probe Applications

Methods for the use of the imaging system of the electron microscope for the production of direct images, crystalline diffraction patterns and small angle reflexion images have been described. A further series of methods, making use of very fine electron probes (finely focused electron beams), have been developed. In principle, the electron optical system of the microscope can be used to produce an electron probe, comparable in diameter with the resolution

of the instrument, simply by reversing the electron optical system so that it is used to produce a demagnified image of the source, the final lens being designed to give minimum spherical aberration, as for the normal objective lens in the transmission instrument.

There are two variants of the probe techniques. In the first the specimen is placed a short distance away from the focused probe, which is used as a fine source, either to throw a projected image or as a means of giving a high degree of coherence for diffraction. In the second the specimen is placed coincident with the probe and one or more of the aforementioned interactions measured from point to point across the specimen, usually by scanning the beam across the specimen in a television-like raster. In the latter case, the final information can be displayed in a two-dimensional picture on a cathode ray tube, for instance.

A brief description will now be given of each of the various probe techniques with some comments on their application and usefulness.

Point Projection (or Shadow) Electron Microscope

By placing an object a short distance after the focused probe, and fluorescent screen or photographic plate a large distance further away, a point projected image is obtained (Boersch, 1940). In principle, the resolution of this image is limited mainly by the probe size and this is subject to the same limitations, due to spherical aberration and diffraction, as that of the transmission electron microscope. Thus, a resolution comparable with that of the transmission instrument should be obtained. One advantage of this technique arises from the fact that the beam does not pass through an electron lens subsequent to the specimen so that there is no loss of resolution or contrast due to inelastic scattering and chromatic aberration. On the other hand, the contrast cannot be so precisely controlled by an aperture as it can in the transmission instrument and the field of view is somewhat limited by the spherical aberration of the final imaging lens (Haine and Mulvey, 1952).

The Point Projection X-ray Microscope

A modification of the point projection electron microscope is achieved by placing a thin metal target in the focused electron probe to give a fine source of X-rays, in front of which a specimen may be placed to give a point projected X-ray image. Several advantages result from such an arrangement, especially as the thin X-ray target can be arranged as a vacuum window, so that the specimen can be placed in air, and very much thicker specimens can be used which would then be transparent to electrons.

This method was originally suggested by Zermak (1896) who proposed the use of a pinhole to give the fine source. Later Sievert (1936) further considered the possibilities and von Ardenne (1939) proposed the use of electron lenses to produce a fine electron probe to determine the X-ray source size. Considerable development on the method has been carried out at the Cavendish Laboratory, Cambridge (England), by Cosslett and his co-workers. The instrument described by Cosslett and Nixon (1953) has two magnetic lenses and a conventional electron gun, giving an electron probe between 0·1 and 1 micron diameter. A very thin metal target window is used with the specimen about 0·01 cm away. A beam current of up to 1 microampere is obtained at an accelerating voltage of 5–20 kV.

The resolving power of the method is limited mainly by electron diffusion in the target, but intensity considerations render the technique increasingly difficult as efforts to improve resolving power increase. Thus to obtain a smaller source, the target must be thinned to reduce sideways diffusion of electrons, reducing also the electron target interaction and X-ray intensity. The smaller electron probe gives a smaller current ($\propto 1/r^2$) for a constant current density, but to produce a smaller spot the beam angle must be reduced, to offset the spherical aberration and, hence, the current density is also reduced. All these factors lead to longer exposure times. The use of an image intensifier may mitigate the difficulties somewhat.

The method is strongly competitive with the micro-

radiographic one, in which a specimen is placed in close contact with a high resolution photographic plate and irradiated with an approximately parallel beam of X-rays. The resolving power is now limited by the plate resolution or, if the latter is small enough, by the resolving power of the microscope which is used to observe the plate. Resolving powers so far obtained are in the region of 0·5 microns.

An excellent review of the methods and applications of the X-ray microscope techniques is given in the Proceedings of a Symposium on the subject edited by Cosslett, Engström and Pattee (1957).

Scanning Techniques

In the second method of applying the electron probe, it is scanned over the specimen to explore its properties point by point. The scanning is normally applied in a two-dimensional raster such as is used in television. The electron beam is, of course, easily deflected to give such a raster, by conventional electrostatic deflector plates or magnetic coils. The probe is focused directly on to the specimen, various properties of which can be measured, point by point, by appropriate detector arrangements.

A transmission specimen may be followed by an aperture which will pass only those electrons falling within the original angle of the probe beam. The transmitted electron current can be measured, and an amplified signal applied to a cathode ray tube scanned synchronously with the probe. Thus a picture is obtained whose contrast depends upon the fraction of electrons scattered outside the contrast aperture by the specimen. Alternatively, the unscattered beam can be blanked off and a dark field image obtained.

The scanning electron microscope was first proposed by Knoll (1935). Later on an instrument was built by von Ardenne (1938). At about the same time an instrument was built by Zworykin, Hillier and Snyder (1942).

There has been no indication so far that the scanning technique used for transmission specimens offers any special advantage. Though it would appear to promise the

possibilities of better control of contrast, and the absence of the effect of chromatic aberration of the imaging lenses, nevertheless it suffers from the very grave disadvantage that every point must be exposed independently in a time sequential operation. Since the probe intensity cannot be greater than the illumination intensity in the transmission instrument, for the same resolving power, it being subject to exactly the same limitations, andsince only one resolved area is examined at a time, the total exposure time must be longer than that in the transmission instrument by a factor equal to the number of image points in the field of view. This factor can be partly offset by the very high sensitivity of electron multipliers, for example, in detecting the final current. However sensitive such detectors might be, they can offer little real advantage in an instrument working near the limit of resolution of the transmission instrument, since this is already limited by quantum considerations which cannot be offset by increased detector efficiency.

Greater advantage for the scanning technique promises when it is used with bulk objects which, for one reason or another, cannot be satisfactorily studied by the replica technique. McMullan (1953) has developed the scanning microscope for the examination of solid surfaces. His instruments use electrostatic lenses and scattered electrons are collected in an electron multiplier to give a high amplification before display. He shows that the back-scattered electron intensity varies rapidly with the angle of illumination particularly between 20° and 30°. This effect is an important source of contrast which will vary as the elevation angle of the surface topography. Resolution is limited by probe size and electron penetration, the latter limitation depends on several factors such as the atomic weight of the object, geometry and irradiating electron energy. A limit of about 100 Å might be achievable, but the best results so far show little better resolution than 0·5 microns. The limitations of intensity and exposure time would become quite severe for a resolution of 100 Å. Some interesting applications have been described (Smith and Oatley, 1955).

The X-ray Probe Microanalyser

The above method promises some usefulness in very specialised applications, but it is unlikely to reach the future importance of the X-ray probe microanalyser originally devised by Castaing and Guinier (1949), Castaing (1954). These workers focused a fine electron probe ($\sim$1 micron) on to a solid surface and passed the X-rays so produced through an X-ray spectrometer. In this way a quantitative elemental analysis of the area of the solid irradiated can be obtained, if the intensities of the characteristic X-radiation lines are measured. Calibration is necessary, using standard elemental or compound specimens, and corrections must be made for absorption and fluorescent excited radiation. Background white radiation also complicates the analysis. Nevertheless, an accuracy of a few per cent can be obtained and elements present in concentrations of less than 1% can be detected.

Resolving power is again limited by electron penetration and diffusion. Wittry (1958) has analysed these limitations. The results depend on atomic number and other factors. Under some conditions a resolving power as low as a few hundred Å might be expected, but 1 micron would be more usual.

The efficiency of production of characteristic X-rays falls off rather rapidly for lower atomic number elements. Also, difficulties of diffracting and detecting the longer wavelength radiation increase. For these reasons elements lighter than magnesium cannot easily be analysed. However, the use of pulse height analysis techniques may provide considerable help in this direction.

The method as above described gives spot measurements. By moving the object the distribution of an element across a specimen (e.g. across a metallic grain) can be followed. It is a natural development to add a scanning facility so that a two-dimensional picture can be obtained of the distribution of any element present (Cosslett and Duncumb, 1957). Even without the somewhat tedious corrections necessary to obtain a reasonable quantitative analysis, the method has

many obvious applications as a semi-quantitative tool and a rapidly expanding interest is being shown in it. Already one commercial instrument is available in France and two further models have been announced in Britain (Duncumb and Melford, 1960; Mulvey, 1960). Active progress is being made towards extending the range of application to the lighter elements.

Other Techniques

The related techniques above described are the ones which, at the present time, offer most hope of widespread practical application. There are, however, still further developments such as emission microscopes, ion microscopes and point emission microscopes which have considerable interest, but which are not discussed, either because their applications are not really very close to those of the transmission electron microscope, or because they are still in so early a stage of development that their useful application cannot yet be foreseen.

REFERENCES

VON ARDENNE, M. *Z. Phys.*, **109** (1938), 553.

VON ARDENNE, M. *Z. Teck. Phys.*, **19** (1938), 407.

VON ARDENNE, M. *Naturwiss.*, **27** (1939), 485.

BOERSCH, H. *Jahrbk.*, A.E.G. *Forsch. Sonderh. Uber mikroskop*, **7** (1940), 34.

VON BORRIES, B. and JANZEN, S. Z. *Verein. Dtsch. Ingen.* (V.D.I.), **85** (1941), 207.

CASTAING, R. *Proc. Internat. Conf. on Electron Microscopy*, London (1954), 60. Published by the Royal Microscopical Society.

CASTAING, R. and GUINIER, A. *Proc. Conf. Electron Microscopy*, Delft (1949), 60.

CERMAK, P. *Ann. Phys.*, Lpz., **60** (1896), 760.

COSSLETT, V. E. and NIXON, W. C. *J. Appl. Phys.*, **24** (1953), 616.

COSSLETT, V. E., ENGSTRÖM, A. and PATTEE, H. H. *X-Ray Microscopy and Microradiography* (Symposium Proceedings, Cambridge, 1956), Academic Press, New York (1957).

COSSLETT, V. E. and DUNCUMB, P. *X-Ray Microscopy and Radiography* Academic Press, New York (1957), 374.

DUNCUMB, P. and MELFORD, D. A. *X-ray Microscopy and X-ray Microanalysis.* Elsevier, Amsterdam (1960), 358.

FERT, C. and RAPORTE, R. *C.R. Acad. Sci.* (Paris), **235** (1952), 1490.

FERT, C. *C.R. Acad. Sci.* (Paris), **238** (1954), 333.

FERT, C., MARTY, B. and RAPORTE, R. *C.R. Acad. Sci.* (Paris), **240** (1955), 1975.

FERT, C. *C.R. Acad. Sci.* (Paris), **241** (1955), 1456.

HAINE, M. E. and HIRST, W. *Brit. J. Appl. Phys.*, **4** (1953), 239.

HAINE, M. E. and MULVEY, T. *J. Opt. Soc. Am.*, **42** (1952), 763.

HAINE, M. E., PAGE, R. S. and GARFITT, R. G. *J. Appl. Phys.*, **21** (1950), 173.

KNOLL, M. *Z. Tech. Phys.*, **16** (1935), 767.

KUSHNIR, YU. M., BIBERMAN, L. M. and LERKIN, N. P. *Bull. Acad. Sci., U.S.S.R. Ser. Phys.*, **15** (1951), 306.

MCMULLAN, D. *J.I.E.E.*, Pt. 2, **100** (1953), 245.

MENTER, J. W. *J. Inst. Metals*, **81** (1952), 163.

MENTER, J. W. *Phil. Mag.*, **44** (1953), 1408.

MULVEY, T. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958.* Springer, Berlin (1960), 68 and 263.

RUSKA, E. *Z. Phys.*, **83** (1933), 492.

SMITH, K. C. A. and OATLEY, C. W. *Brit. J. Appl. Phys.*, **6** (1955), 391.

SIEVERT, R. *Acta. Radiol.*, Stockholm, **17** (1936), 299.

WITTRY, D. B. *J. Appl. Phys.*, **29** (1958), 1543.

ZWORYKIN, V. K., HILLIER, J. and SNYDER, R. L. *Bull. Am. Soc. for Testing Materials*, **117** (1942), 15.

CHAPTER X

SPECIMEN TECHNIQUES AND APPLICATIONS

Introduction

THE range of application of the electron microscope has become so wide that it can only be discussed in purely general terms in this single chapter. It was thought earlier that applications would be limited to inorganic particulate objects and that biological specimens would give little contrast because of their low atomic number and uniform density, and because they would be destroyed in the electron beam. This was before it was realised how important phase effects could be, and before the modern refinements of specimen preparation were appreciated. In fact, wide and spectacular successes have been achieved in the field of biology, especially since methods of cutting ultra-thin sections have been perfected. Metallurgical applications have been less spectacular but, nonetheless, extremely important, particularly since modern replica techniques and methods of preparing very thin films of metals have been developed. Indeed, the electron microscope is now finding uses in almost all branches of science and technology.

Types of Structure

Structures which are met in microscopy may be divided into five general categories:

1. *Ultra-microscopically particulate*, consisting of very fine particles, each of a relatively homogeneous nature and usually of interest in regard to their size and shape distribution but sometimes possibly for their crystalline properties.

2. *Microscopically particulate*, comprising particles of matter which can be observed directly in the microscope and having internal structure and shape which is of direct interest. This category would include the minute living organisms, bacteria, for example.

3. *Surfaces*, the topography of which is often of very great importance, particularly in metallurgy.

4. *Bulk, homogeneous structure*. This is a structure which, though generally of a random nature, is repeated fairly uniformly throughout the bulk of a material. An example is the type of structure which is met in a polycrystalline solid such as a metal, and in some types of biological material. In general this type of structure can be fully explored by means of a single cross-section since this would be representative of the material as a whole.

5. *Bulk, heterogeneous structure* which, though showing local regularity, may vary greatly from one part of a body to another. This is typified by most examples of biological material. A random sample of such material can only give very limited information. Comprehensive sampling throughout a body may be necessary to build up a complete picture.

It would clearly not be possible to classify all material under these headings, nor would it be possible to give a specific range of techniques for each type. Indeed, any one material may overlap to a large extent the various headings given. For example, a body may have a surface whose topography is of interest, it may have a bulk heterogeneous structure of, say, a fibrous nature on a macroscopic scale and the fibres may all have the same homogeneous microscopic structure. To obtain the fullest possible information on such a body a combination of several optical and electron optical specimen preparation methods may be needed, as well as X-ray, electron diffraction, absorption spectroscopy and chemical analysis. It will be our limited object here to outline the more important specimen techniques of use with the electron microscope and to give examples of their applications. Some comments will also be made as to possible extensions of the techniques and, wherever possible, reference will be made to relevant

published literature to guide the specialist beginner in his further reading. Most of the basic methods are described in *The Practice of Electron Microscopy* (Drummond, 1950). Developments since that time are usually to be found briefly recorded in the biennial review articles in *Analytical Chemistry* (Swerdlow, 1954; Swerdlow, Dalton and Birks, 1956; Rochow and Botty, 1958; Rochow, Thomas and Botty, 1960). A shorter review of preparative methods has been given by Cosslett and Horne (1957). Methods for biological material are fully covered by T. F. Anderson (micro-organisms) and Sjöstrand (cells and tissues) in Vol. III of *Physical Techniques in Biological Research* (1956) and in *The Microtomist's Vade-Mecum* (Lacy, 1960). The most recent developments will be found in the *Proceedings of the Fourth International Conference on Electron Microscopy*, Berlin, 1958 (Springer, Berlin, 1960) and in a recent German text (Reimer, 1959).

Specimen Support Films

It must always be borne in mind that matter is never entirely transparent to electrons. For this reason the electron microscope specimen requirements differ from those for the optical microscope. To achieve best contrast and resolution all superfluous atoms should be removed where possible. In the optical microscope it is possible to resolve a structure near the resolution limit of the microscope even when embedded in or mounted on a transparent material 10,000 times thicker than the resolving power. In the electron microscope the maximum value of this ratio is about 10, if serious loss of contrast is to be avoided. This leads to a good guiding principle in specimen preparation: 'keep the amount of material in the object down to a minimum', including the thickness of any supporting film. It is probably true to say that the wide variation in quality of electromicrographs produced by one operator and another, and one laboratory and another, is much more affected by this axiom than by the variation in quality of the instruments used.

Just as in the optical microscope it is, in general, necessary

to supply a glass slide to support the specimen, so in the electron microscope an equivalent support must be provided. As has been said above, this support must be extremely thin and therefore relatively strong. For high resolution working this support must not exceed a few tens of Å in thickness, and yet it must be itself stretched over a supporting frame giving a reasonable field of view. In the early microscopes such supporting films were stretched over a single circular aperture in a platinum disc, the aperture diameter being 25–50 microns. This arrangement gave a very restricted field over which to search for interesting details and was soon replaced by a grid structure, an arrangement which has been generally used since its first inception (Figure 10.1). The grids in most common use

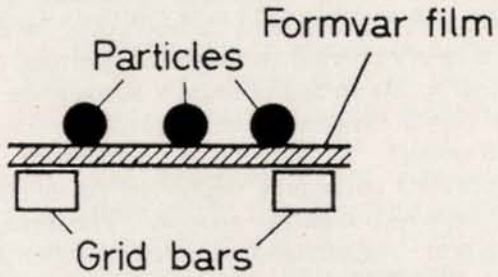


FIG. 10.1. *Specimen resting on support film, carried on metal grid (Bradley).*

are of copper, about 3 millimetres in diameter, 25–50 microns thick and contain a square mesh (200 per inch) with windows 90 microns square and 35 micron bars. A variety of patterns are now available for special requirements, e.g. with a set of parallel bars, without cross bars, or for mounting very thin sections.

The use of copper in the specimen support grid is important to give the best possible thermal conductivity. An improvement in this respect could probably be made by the use of thicker grids; it has not, however, been possible to obtain grids with the bars more than a certain thickness since the photomechanical method used to manufacture them limits this dimension.

The requirements for the specimen support film itself are that it should be extremely thin, preferably less than 100 Å, it should be amorphous and structureless, it should be strong and capable of withstanding the effects of the electron beam, and its preparation should be relatively simple. For a long time cellulose nitrate films were in common use. These were prepared by spreading a dilute solution of the nitrate in a suitable solvent on to a water surface, allowing the solvent to evaporate and then collecting the film on a grid by raising the grid on a suitable holder up through the water to the film. To a certain extent this type of film was superseded by the use of thin beryllium or silica foils, produced by evaporation on to a suitable substrate. But all such films have now been replaced by the evaporated carbon film, originally proposed by Bradley (1954; 1956), who showed that carbon could be evaporated by passing a heavy current through two pointed carbon arc rods with the points in light contact. Resistance heating at the point of contact raises the temperature to a value where evaporation occurs.

The evaporated carbon is deposited on a structureless substrate from which it can be stripped. The apparatus used for the carbon evaporation is shown schematically in Figure 10.2. Two pointed carbon rods are arranged in a vacuum bell jar, one fixed and the other sliding in a silica tube and lightly sprung so as to force the two points lightly

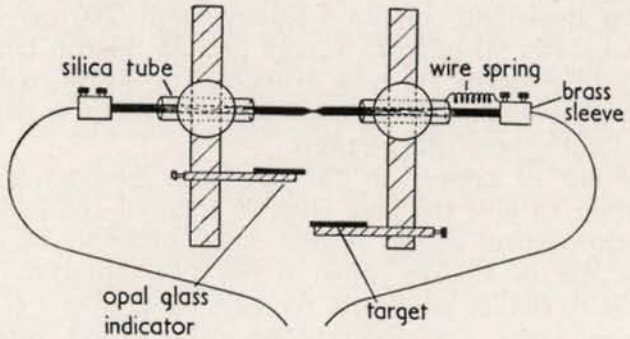


FIG. 10.2. *Carbon evaporation plant, schematic.*

together. An alternating current, of between 20 and 50 amperes, from a small transformer is passed through the rods and results in intense heating in the region of contact and evaporation on to the target platform. It is desirable first to degas the electrodes by raising the current to just below the value for evaporation once or twice, as otherwise large fragments may be thrown off the electrodes during the final evaporation.

Bradley describes a simple method of estimating the evaporated film thickness. A small piece of white porcelain is mounted below the point of evaporation and on it is placed a small drop of low vapour pressure vacuum oil. The carbon evaporates on to the porcelain and thus darkens it except in the region of the oil drop where, for some reason, no darkening occurs. He found that when the darkening was just visible the thickness of the film on the indicator was about 100 Å with the indicator placed about 8 centimetres from the source. Thinner films could be obtained, while still using the indicator in the region where reasonable darkening occurred, by placing the target at a larger distance away. The practicability of determining more exactly the thickness of the film by measuring its optical thickness has been investigated by Agar (1956). The method is only reliable, however, if the conditions of evaporation are carefully standardised (Cosslett and Cosslett, 1957).

The carbon is evaporated on to a substrate which can easily be stripped from a glass slide, or directly on to the slide itself or on to mica. Suitable substrates are Bedacryl 122X, Boron oxide or Glycerol; Bradley recommended Bedacryl*. To coat the target slide the Bedacryl is first diluted with redistilled Benzene to a strength of 6–8% weight/volume. A quantity is then poured over a clean microscope slide and allowed to drain off. The film dries in a few seconds under a lamp. After evaporation the composite film, having been scored into small squares, is floated off on to a water surface. Stripping is facilitated by

* Bedacryl 122X, as supplied (in 40% solution in xylene) by I.C.I. Ltd., England.

breathing heavily on to the film. The floating film is next collected on to a supporting grid after which the Bedacryl supporting film is washed off with a mixture of 50% ether and 50% acetone. Full details of this method of preparation have been given by Bradley (1954).

Alternatively the carbon may be deposited directly on to the surface of freshly cleaved mica. In this way it is possible to obtain films of thickness 50–60 Å and of a few sq. cm in area, so that several grids can be covered at one time. It is also often possible to strip a carbon film which has been deposited directly on a glass slide, without substrate, but much depends on the state of its surface. A slide which has been used for other purposes should first be cleaned with a detergent and then lightly polished with lens tissue; with a new slide it is usually sufficient to polish after lightly breathing on it.

Evaporated carbon films have several advantages over the type of film previously used. In the first place their production is relatively very simple and is greatly helped by the fact that the films remain visible to the eye down to much lower thicknesses than is the case with the other known films. This factor is extremely important; silica films become quite invisible at the small thicknesses required and have to be handled with a good deal of 'faith'. The carbon films are almost amorphous, extremely strong and withstand quite high temperatures. They can be supported across a 200-mesh grid even when their thickness is less than 50 Å. They are also chemically inactive and have a relatively high electrical conductivity, so that they are not affected by electron bombardment. The wide use of evaporated carbon in replica techniques will be discussed later.

The Mounting of Specimens on Support Films

Specimens may be mounted on support films by a number of methods, including direct manipulation, settling from air, flotation, settling from a liquid suspension, or vacuum evaporation.

The most generally used method is that of deposition

from a drop of liquid on to a filmed grid. In many cases the end result of separating material for examination (by filtration, centrifugation, etc.) will be a liquid suspension, as with viruses, bacteria, colloids. If the specimen is in the form of a dry powder, it may be dispersed in a suitable medium by shaking or by rubbing out with a spatula or muller. Details of special methods are given by Drummond (1950). The suspension may be transferred to the grid with a glass rod or a micro-pipette; trial will soon show what degree of dilution will give a reasonably covered field of view in the electron microscope. Alternatively, the suspension may be sprayed on to filmed grids with an airblast (atomiser); details of an apparatus suitable for viruses is given by Backus and Williams (1950) and refinements in its use are described by Nixon and Fisher (1958). Certain crystalline materials may be allowed to form microcrystals directly on a support film (Hall, 1950a; Dawson, 1952).

Dry powders may be dusted directly on to a grid, if it is undesirable to put them into liquid suspension. When they have to be collected from an aerosol, or from a flame reaction, a special type of thermal precipitator may be used (Drummond, 1950; Schlipköter *et al.* 1960). Special methods have also been devised in a number of technological fields, to meet the need for rapid and reproducible preparation of inorganic particles e.g., for clays, carbon blacks, colloids, paints (cf. Swerdlow (1954); Swerdlow *et al.* (1956)). Some materials may be examined in the form of thin films, formed by electrolytic deposition or by vacuum evaporation, particularly when the process of formation itself is being investigated (cf. Holland, 1956).

Shadow Casting

Many specimens are found to be too thin to show adequate contrast in the electron image; this is especially true of very small biological fibrils or particles, such as plant viruses. Visibility of the preparation can be greatly enhanced by depositing a layer of metal obliquely on to it, by evaporation *in vacuo*. The specimen, which may already be mounted on a supporting film and grid, is

placed at a distance of 15–20 cm from the evaporating source, which may be a V-shaped filament or a tantalum boat, in a bell jar (Figure 10.3). A 'shadow', free of metal,

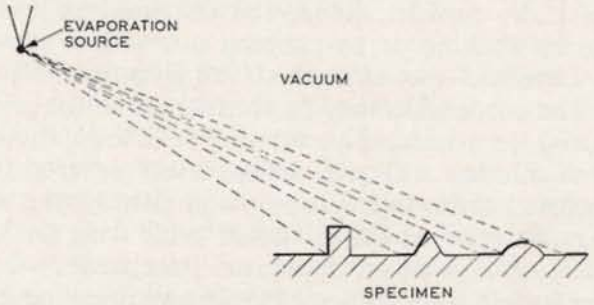
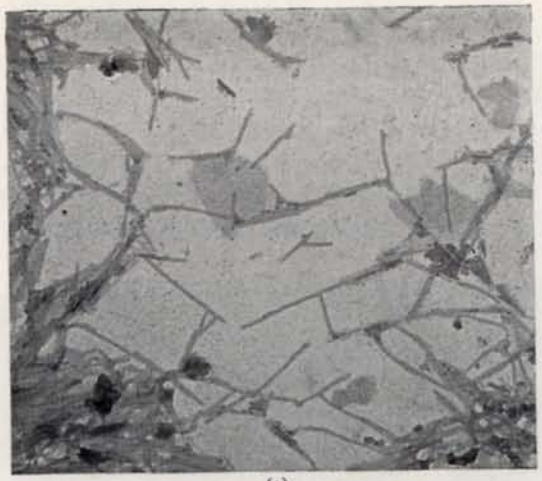


FIG. 10.3. *Arrangement for shadow casting: metal is evaporated on to the specimen, at an angle, from a hot filament in a vacuum bell jar.*

will be formed on the lee-side of any features which project above the general level of the support film, whilst the front side will collect a correspondingly greater amount of metal. Owing to the high scattering power (particularly of the heavier metals) for electrons, these features will be outlined with great clarity in the image. A shadow-cast specimen, when photographed, is usually printed or projected as a negative; the shadows (free of shadowing material) then appear dark and the regions covered with the shadowing material light. The image appears as if illuminated obliquely, i.e. as if the shadowing atoms were a beam of illumination. In this way a 'picture' of familiar type is obtained leading to easy interpretation (see Figures 10.4 (a) and (b)). If the angle of deposition is known then the height of specimen details can be estimated from the length of the shadows they cast.

A number of types of commercial evaporator are now available, the essentials of which are a fast pumping system and a large bell jar into which stout current leads have been sealed. It is helpful to build in an aperture, or set of apertures, so that the metal beam reaches the specimen from a well-defined direction only, and not by reflection from the



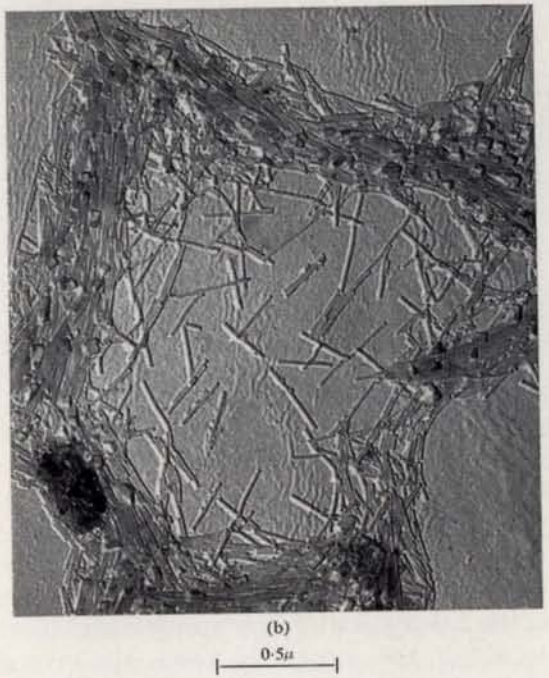


FIG. 10.4. *Tobacco mosaic virus,* (a) *unshadowed, and* (b) *shadowed with chromium* (*R. W. Horne, Cavendish Laboratory*).

walls of the apparatus; this is especially important if carbon is to be evaporated (see below). A photograph of a simple evaporator, which is readily put together in a good workshop, is shown in Figure 10.5. The vessel is a metal cylinder, instead of a bell jar, and evaporation takes place upwards, so that the specimen may be cooled if desired.

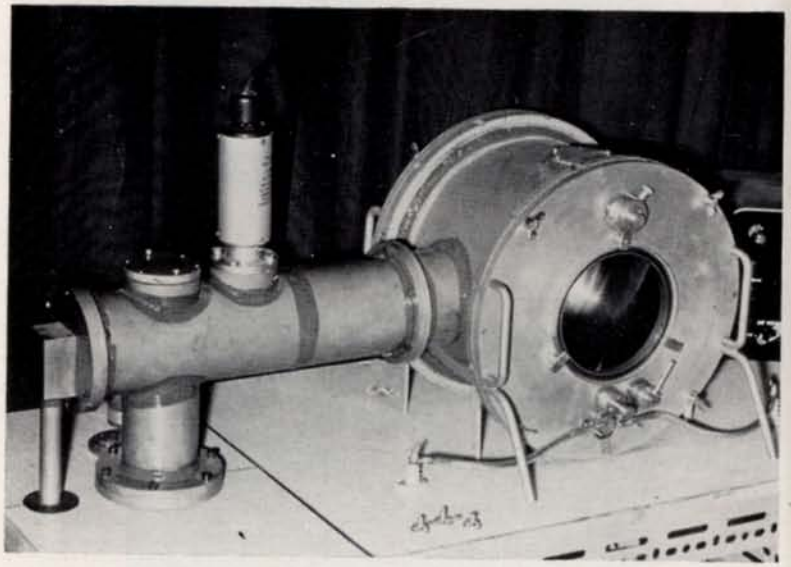


FIG. 10.5. *Shadow-casting plant, with provision for apertures and a liquid air trap for cooling the specimen (R. W. Horne and H. E. Pearson, Cavendish Laboratory).*

It is important not to evaporate too much metal or important details of the specimen may be obscured. A mono-layer of the heavy metals, such as gold, platinum, uranium, will give adequate contrast, but several mono-layers of a lighter metal such as chromium will be needed. Drummond (1950) discusses the requirements in detail. In making the calculation of how much metal to evaporate in order to obtain the chosen thickness, the inverse square law may be assumed to be valid if an aperture is used between source and specimen. In high resolution microscopy

it is found that almost all metals aggregate too much, and show a grainy surface which might be confused with object detail. This defect is accentuated if the specimen is over-heated in the electron beam. A mixture of gold-palladium gives better results, and wire of a suitable alloy is commercially available. Even this shows grain of order 25 Å, and beyond this resolution only the carbon-platinum mixture devised by Bradley (1958, 1959) appears to be adequately structureless. The material may be evaporated from a composite rod, as described by him, or a small piece of platinum wire may be twisted around the point of contact of two carbon rods, as used in the normal method of preparing carbon films. In this method, even more than when evaporating metals, it is necessary to site the specimen far enough away from the source (15–20 cm) to diminish the rise in temperature. The shadows are then extremely sharp with the platinum-carbon mixture (Figure 10.6).

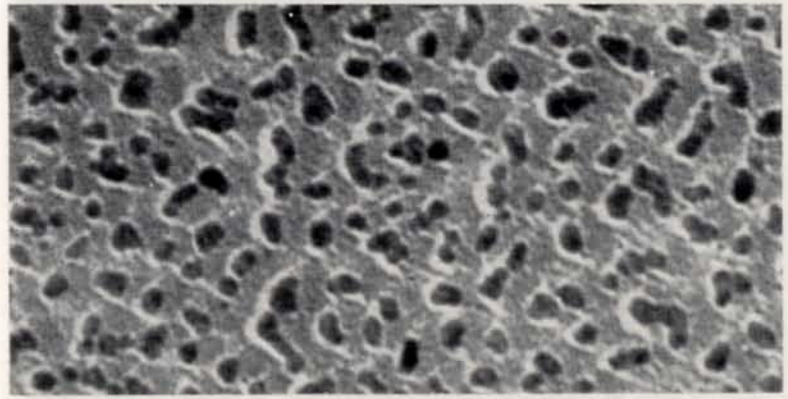


FIG. 10.6. *Shadows formed by evaporating a platinum-carbon mixture are extremely sharp.*

Main Types of Specimen

Thin Films

There are a number of types of specimens which occur in the form of thin films and can be mounted directly on specimen grids for examination. An example of this is a thin flake of mica, broken from a mica plate.

Evaporated films are commonly used for many purposes, for example for the production of electrical resistors, conducting coatings, optical films, and in several types of special image tube. The continuity and crystalline structure of these films is often of importance and, if not too thick, they may be examined in the transmission electron microscope (provided they can be separated from any thick supporting substrate, on which they may be evaporated, for insertion in the microscope). Very often they can be evaporated directly on to carbon films for examination, though it must always be borne in mind that the substrate can play an all-important part in determining the final structure, so that a film evaporated on to carbon may be quite different from one evaporated on to, for example, a glass surface (Holland, 1956).

In a number of metallurgical investigations, it is desirable to prepare samples which will be thin enough for electron microscopy and yet still be representative of the bulk material. Although attempts have been made to cut thin sections, in the same way as for biological tissues, the results have not been very satisfactory owing to the distortion and possible work-hardening introduced in the cutting process. Castaing (1956) and Fert, Marty and Saporte (1955) developed a method of ion bombardment which is very effective, but not very easy to control as regards final film thickness. Its main value would appear to be for the removal of successive surface layers from a metal, during observation in the reflexion electron microscope. A more easily regulated reduction in thickness may be produced by etching with solvents or electrolytically. Heidenreich (1949) and later Hirsch, Horne and Whelan (1956) obtained good results with aluminium and its alloys, and a method suitable for ferrous metals has been described by Brandon and Nutting (1959). A detailed account of preparative methods has been given by Kelly and Nutting (1959). With practice, it is possible to get films of any desired thickness down to about 100 Å. They have been much used in the study of dislocations and other imperfections of structure (Figure 10.7); for a survey of this

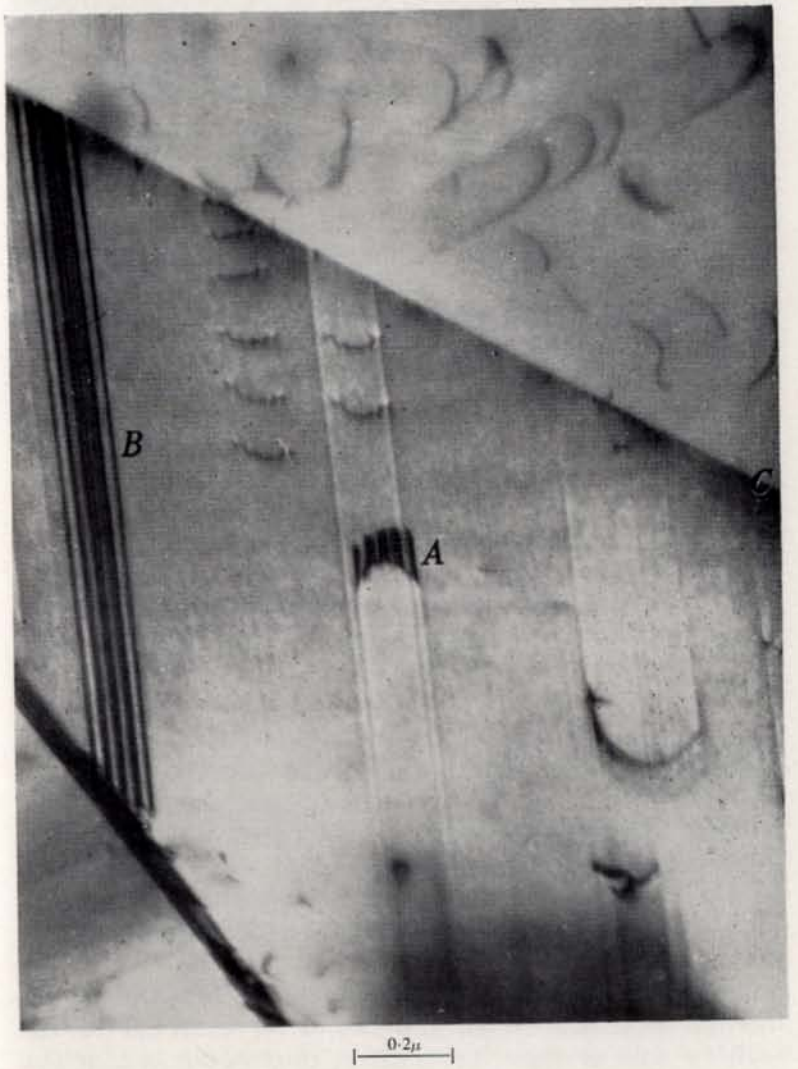


FIG. 10.7. *Dislocations and extended stacking faults in thin film of stainless steel; transmission micrograph (M. J. Whelan, P. B. Hirsch, R. W. Horne and W. Bollman, Cavendish Laboratory).*

field see Hirsch (1959) and the recent Symposium arranged by the Institute of Metals (*J. Inst. Metals*, August, 1959).

Fibres

Fibres may also fall into the category of self-supporting objects, but they are, very often, excessively thick and allow little really useful information to be obtained by such direct observation.

For the study of the internal structure of fibres, the thin sectioning method described below gives most information. However, the method of packing of the sub-fibrils is also of interest and this may be studied by controlled disintegration. Methods of mechanical, sonic and supersonic disintegration have been developed for this purpose, as well as chemical degradation with acids and alkalis. Details of the procedures are given by Drummond (1950), Sjöstrand (1956) and Sikorski (1960). They have been applied particularly in the study of cellulose fibres, but also for hemp, jute and silk, as well as for man-made fibres.

Methods for the examination of fibre surfaces by the reflexion technique have also been developed and are briefly described in a later section.

Thin Sections

Almost all biological tissues are too thick for direct observation in the electron microscope. In some instances it is possible to isolate components of special interest, such as plant cell walls (Green, P. and Chapman, G., 1955) or mitochondria (Anderson, N. G., 1956) by chemical action or mechanical disintegration (cf. Drummond, 1950, p. 85), but such action is usually too drastic. In order to investigate the microstructure of tissue, especially the relative situation and inter-connection of the component cells and of the internal organisation of the cell, it is essential to cut sections thin enough for electron microscopy, i.e. a few hundred Å thick. The microtome, as used for preparing sections for optical microscopy, will rarely cut sections thinner than 1 micron. In recent years, however, its mechanism has been improved to such a degree that

sections as thin as 200–300 Å can now be cut in routine fashion (Figure 10.8). 'Ultra-microtomes' of a variety of designs are now commercially available, and details of their construction will not be given here. They differ from the instrument used for cutting sections for optical microscopy essentially in the much slower rate of advance, or 'feed', of the sample towards the knife, which is rigidly fixed to the base plate. The advance may be either thermal, by electrical heating of a long bar, or mechanical, by means of a screw or gear system. The latter method is more easily variable, but has to be very carefully designed if backlash is to be eliminated. Whichever system is used, the instrument as a whole must usually be shielded from draughts since a very small rise in temperature may easily cause an advance or retardation comparable with that produced by the mechanism.

Details of instruments and techniques are given fully by Sjöstrand (1956) and by Lacy (1960); the physical processes involved in sectioning are discussed by Sitte (1960). The most difficult steps are the embedding process and the collection of the cut sections. It is now usual to use a glass knife cut from a sheet of plate glass, but specially sharpened diamonds have also been used and some workers still favour a steel blade. A miniature water bath is fixed to the back of the cutting edge, so that the sections are drawn off on to the surface of a water-alcohol or water-acetone mixture as they are cut. They are fished out on a specimen grid, under a binocular microscope in a properly directed beam of light. Patience and practice are needed, especially to obtain ribbon sections.

The methods of fixing and embedding biological material are still in process of development, and opinions differ widely on the best procedure for each particular type of specimen. It seems probable that no single fixative, or embedding medium, is universally applicable and that each type of tissue will require separate consideration, depending on its chemical composition and physical structure. The polymerisation of methacrylate in the specimen, the first successful method of embedding for ultra-microtomy,

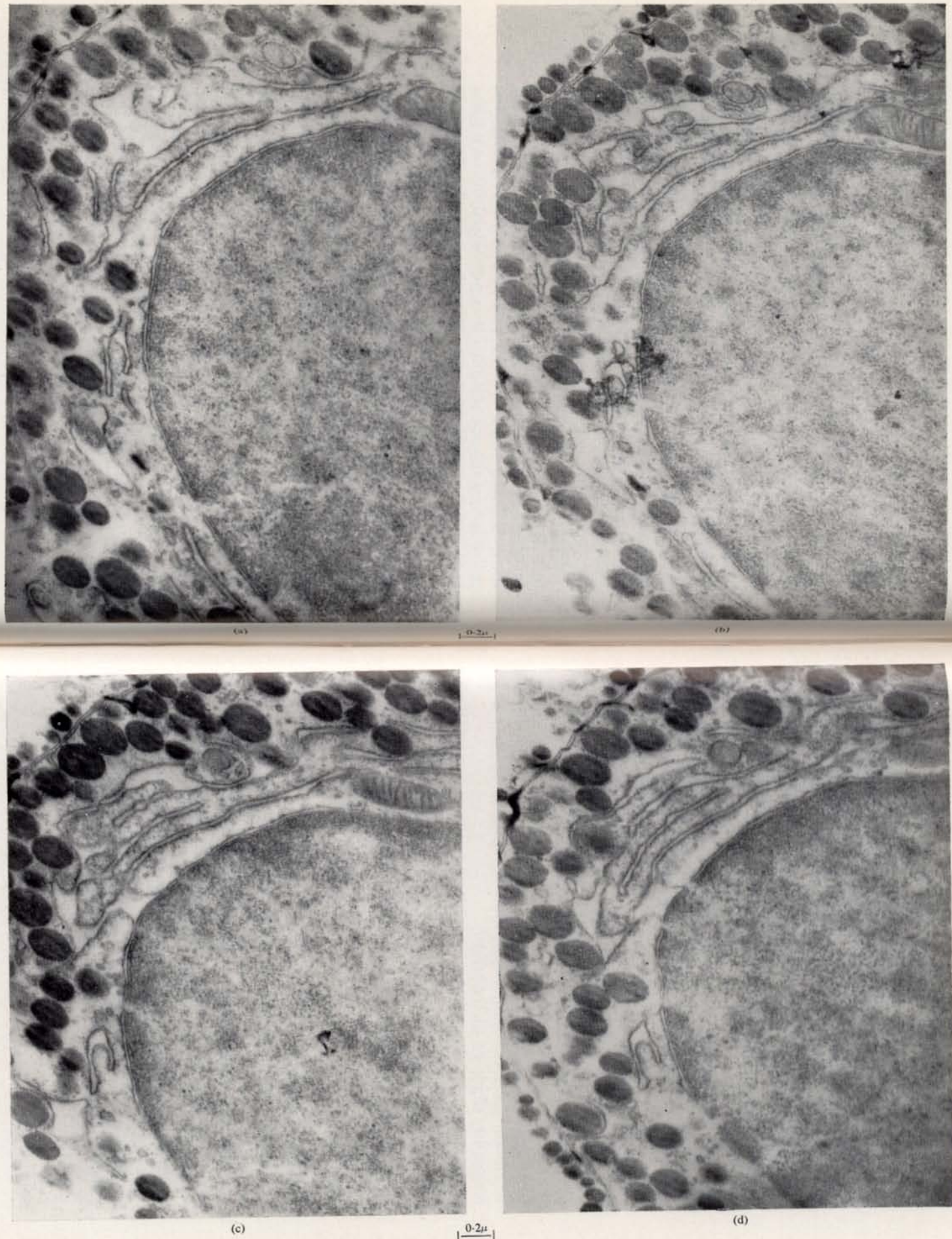


FIG. 10.8. (a, b, c, d) *Serial sections of mouse anterior pituitary, treated with phosphotungstic acid and embedded in araldite* (*Mrs. B. Barnes, Cavendish Laboratory*).

produces artefacts in at least some tissues. The use of an epoxy resin ('Araldite') seems open to less objection (Glauert, Rogers and Glauert, 1956; Glauert and Glauert, 1958), and polyethylene has also given excellent results (Turian and Kellenberger, 1956; Ryter and Kellenberger, 1960). In both cases, however, the generic name covers a range of materials of slightly differing composition and viscosity, and it may even happen that batches identical in the eyes of the manufacturer will have varying properties as an embedding medium. It is as well to adhere closely to the materials and methods specified by the originators of these new procedures, and to examine the resulting sections carefully for signs of artefacts, such as voids, fragmentation or displacement of tissue. It is also as well to check the nominal thickness of sections, as given by the microtome advance, by an optical method (cf. Tolansky, 1948) since estimates made on the basis of interference colours (or the lack of them) have been shown to be usually much too low (Peachey, 1958, 1960; Sitte, 1960). It is very difficult to obtain sections thinner than 200 Å, in spite of claims frequently made to the contrary.

Bulk Specimens

Bulk specimens such as metals cannot be observed directly in the normal electron microscope and resort must be made to one of the various replica techniques described in the next section, or to reflexion electron microscopy (see Chapter IX).

Replica Techniques

Together with the rapid development of the section-cutting technique the parallel improvements in replica methods constitute the main advance in specimen preparation technique for electron microscopy during recent years. At first the replica technique was restricted almost entirely to metallic specimens but more recently it has become apparent that it has wide application to many other types of material. The principle of the replica technique is to transfer the surface *topography* of the solid body to a thin

film so that it may be observed in the microscope by transmission. There are many variants of the replica technique, but generally they produce one of two types of replica, as illustrated in Figures 10.9 (a), (b). The replica in (a) has one surface flat and the other following the topography of the replicated surface; (b) comprises a film of uniform thickness following the surface contours.

The replica type (a) has been most generally used until about 1955, though the second type was well known before that time. Latterly, type (b) has come into very general use, partly because the technique for its production has been

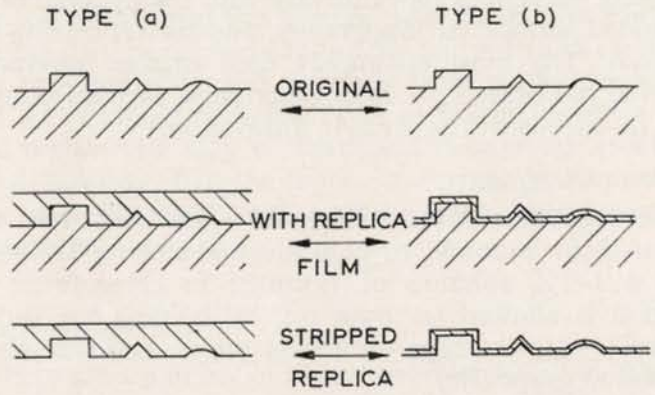


FIG. 10.9. (a, b) *Two types of replicas in replica methods* (a) *and* (b).

greatly simplified and partly because it gives a much improved image. It was at first thought that the improved image resulted from the fact that the (a) type replicas were formed in plastic and were limited by the large molecular size of the replicating material. When, however, Bradley showed that a replica of type (b) taken off the surface of one of type (a) gave a much improved image, it became apparent that the limitation arose from the greater amount of material in the type (a) replica.

The process of making a replica can be divided into three stages:

1. the deposition of the replicating material on the surface to give intimate contact;

2. the removal, or stripping of the replica;
3. the mounting of the replica on the support grid.

It will be seen that these basic stages in the preparation can be extended to a large number of modifications of the basic process. The type (a) replica is produced when a direct solution of a plastic is flowed over the surface, and the solvent allowed to evaporate. The surface tension tends to make the upper surface of the resulting film flat, though, as Agar (1956) has shown, for the thinner films the surface does tend to follow the surface contour to some extent and this in itself leads to loss of resolution. The method of stripping the replica film can vary with the material of the replicated surface, its topography, and the replicating film material. The most commonly used organic replicating material is Formvar. A brief description of the technique used for metals will be given in illustration.

Formvar Replicas

The specimen surface which, it is assumed, has been prepared (for example, by polishing and etching) is flooded with a 1–2% solution of Formvar in chloroform, the solution is allowed to drain off, by holding the surface vertically with the edge on a filter paper, and the solvent allowed to evaporate.

The next stage of stripping can be carried out in a number of different ways; the one described is that used by Bradley in the two-stage carbon replica technique, the second part of which will be described later. To strengthen the film for stripping it may be backed with a layer of another plastic, for example, Bedacryl (see p. 231). This material is applied in a 1–2% solution in Benzene which, after application, is again allowed to drain as for the Formvar. The backing film is allowed to dry thoroughly (2–3 minutes). To strip the replica it is first scored round the edge of the specimen, with a needle or sharp blade, and a length of Sellotape is pressed firmly on to it. The tape is then stripped off, removing the replica with it. Bradley points out the desirability of breathing heavily on to the composite film immediately before applying the Sellotape.

After stripping the composite replica film on the Sellotape the latter is mounted on a glass slide with the replica film outermost, ready for removing the replica from the tape.. The slide is placed in a bath of acetone for about two minutes to dissolve away the Bedacryl backing which is in contact with the tape. The Formvar is insoluble in acetone and remains in light contact with the Sellotape, being anchored round its edges with the adhesive on the tape. The film is now teased away from the Sellotape on one side and a specimen grid slid between the Formvar and the tape, which is then removed from the solvent. The grid is lifted with fine forceps a short distance away from the tape and held until the acetone has dried. The Formvar film can then be cut away round the grid with a sharp blade, thus preventing the wet Formvar from curling under the grid and forming a double layer in the final replica. The final replica film may be shadowed if required. It will be noted, however, that the replica is a negative one and this fact must be taken into account when interpreting the final image.

Type (b) Replicas

There are a number of processes which result in this type of replica: by the evaporation of a film of suitable material on to the surface and its subsequent removal, by the stripping of an (a) type replica on to which is evaporated a thin film of the type (b) replicating material followed by the dissolution of the type (a) replica from it or, in some cases, the replica can be made from a metal surface by anodically oxidising it and then dissolving away the metal to leave the oxide film as the replica (see Drummond, 1950).

The use of silica to produce this type of replica was first proposed by Heidenreich and Peck (1943). The replicas produced give a high resolution image but, owing to the very low visibility of the thin silica films, they are exceedingly difficult to handle and, in fact, since they are largely quite invisible, the process has to be carried out with a good deal of faith. The introduction of the evaporated

carbon technique by Bradley overcame this serious disadvantage as carbon films are easily visible even in very small thicknesses. The carbon films can be evaporated directly on to the surface, but in many cases it is more convenient to use a two-stage process, unless, as is often the case for particulate matter, the specimen can be readily dissolved from the replica.

The extension of the Formvar replica technique, described above, to produce carbon replicas, starts after the composite replica has been stripped from the surface under examination onto Sellotape and mounted on a glass slide with the replica facing outwards. The glass slide is next placed in a bell jar and a carbon film of a suitable thickness (50–200 Å) is evaporated onto it. On removal from the vacuum the Bedacryl layer is dissolved away as before and the composite Formvar/carbon film is mounted on a specimen grid inserted between the Formvar and the Sellotape. The grid is slightly bent before insertion and, after collecting the film, is mounted on a peg of the same size as the grid, with the convex side and the film upward. The grid is first wetted with a drop of acetone, which is allowed to fall upon it from a dropping pipette held about 3 cm above it, and about 3 c.c. of chloroform are subsequently run over the grid at the rate of 1 c.c. per minute from a burette so arranged that its nozzle is 2 mm above the grid mounted on the peg. The chloroform dissolves away the Formvar and leaves the carbon film mounted securely on the grid. The preliminary addition of a drop of acetone avoids the possibility that first contact with the chloroform might pull the grid away from the peg.

The grid is next dried under a lamp, whilst still mounted on the peg, leaving the replica ready for shadow-casting or for examination without shadowing.

Various modifications of this method have been described by Bradley (1954, 1956, 1960). The method just described leaves the negative side of a replica outermost and is therefore not entirely suitable for shadowing. To reverse the film it may be lifted from the Sellotape, on a ring-shaped tool,

FIG. 10.10. (a, b) *Unshadowed and shadowed Formvar replicas and carbon replicas.*

(c)

3μ

(d)

3μ

FIG. 10.10. (c, d) *Unshadowed and shadowed Formvar replicas and carbon replicas.*

after dissolving away the Bedacryl, and then inverted on to the supporting grid.

It will be appreciated that the contrast produced by this type of replica, when unshadowed, results from the effective change in thickness due to the slope of the film relative to the electron beam. Thus the maximum contrast is obtained when the slope is a maximum or, in other words, when the rate of change of topographical altitude is a maximum. The contrast so produced is a sort of differentiated topography and thus produces a picture which is not easily appreciated owing to the unusual nature of the contrast variation. This could be quoted as a disadvantage of this type of replica were it not for the fact that shadowing produces contrast of a very familiar nature, being strikingly similar to that obtained 'in normal life' by oblique illumination. The only deviation which can sometimes confuse, if not borne in mind, is that caused by the superposition of the contrast produced by the shadowing material with that produced by the underlying carbon.

To illustrate the main differences between the different replica techniques Figure 10.10 compares the (a) unshadowed Formvar replica, (b) shadowed Formvar replica, (c) the unshadowed carbon replica and (d) the shadowed carbon replica. The subject is, in each case, a pearlitic steel. The improvement in resolution is very obvious and the improvement in 'pictorial quality', especially for the shadowed carbon replica, is most striking.

A further advantage of the type (b) replica lies in its suitability for stereoscopic viewing. In the type (a) replica, since the contrast is produced by variation in thickness, it is not possible to ascribe a specific topographical altitude to each point on the replica, whereas in the second type such is the case, and striking stereoscopic effects are obtained from two pictures of the same field taken with a specimen tilted at different angles with respect to the electron beam. Figure 10.11 shows a stereoscopic pair from a pearlitic steel specimen.

In respect of resolving power the carbon replica is much superior to the Formvar and silica replicas, as mentioned

earlier. When the evaporation is properly carried out, the carbon film shows no self-structure, at least down to the order of 10 Å. As in preparing carbon support films, however, the precautions mentioned on p. 229 must be

FIG. 10.11. *Stereoscopic pair of replicas of pearlitic steel.*

observed in the course of evaporation if the presence of coarser grains in the deposited film is to be avoided (Cosslett and Cosslett, 1960; Calbick, 1960). The high resolution obtained in a properly prepared carbon replica is illustrated in Figure 10.12.

Application of Replicas

Metals

The widest application of the replica technique must be in the field of metallurgy. The replication process is relatively straightforward here and the two-stage Formvar/carbon technique can most generally be recommended. Usually, etching techniques similar to those used for optical work are sufficient, except that rather lighter etches may be desirable. A detailed appraisal of the value of the replica method in routine metallurgy will be found in the reports prepared by a special sub-committee of the American Society for Testing Materials (A.S.T.M., 1950, 1952, 1954, 1957). On this account, this subject is not discussed further

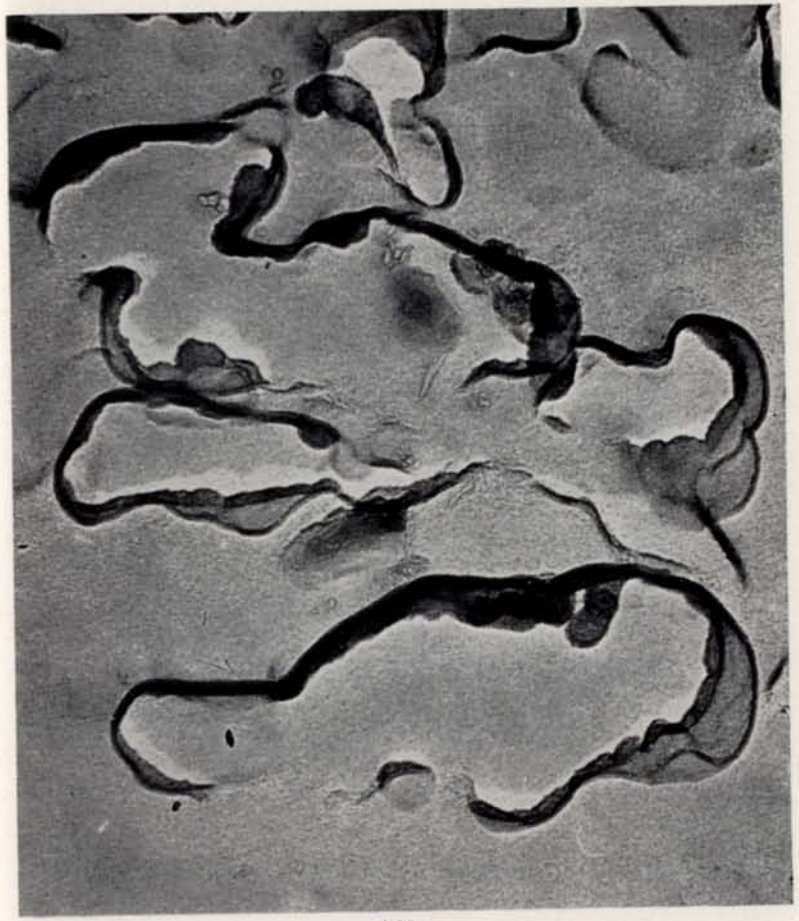


FIG. 10.12. *High resolution carbon replica of tempered sorbitic steel.*

in the present chapter, apart from the extraction replica method (below). With the development of methods of reducing bulk metal to very thin films, noted previously, the use of replicas in metallurgical investigations has somewhat decreased. They have, however, also been found useful in connection with reflexion electron microscopy, when the specimen itself is inaccessible or would be damaged by the electron beam.

Particulate Material

There is a range of particulate material which cannot be examined usefully in the electron microscope directly, because the particles are too large and produce excessive electron scattering. An example of this is shown in Figure 10.13 which shows a group of natural chalk particles as

FIG. 10.13. *Direct micrograph of chalk particles.*

imaged directly in the electron microscope. It is seen that although the general outline is clearly visible no other structural details can be seen.

To obtain a replica of these particles Bradley (1956) deposited them from a distilled water suspension on to a Formvar film mounted on a support grid, and then evaporated on to them a thin film of carbon. The Formvar was dissolved off, leaving the particles supported in the evaporated carbon film. The particles were then dissolved away in a suitable solvent, leaving the carbon replica,

which could be shadowed if required. Figure 10.14 shows the resulting micrograph of typical particles in which the surface detail is brought out very clearly.

This method can be applied to a very wide range of objects, for example, photographic grains (Hamm and

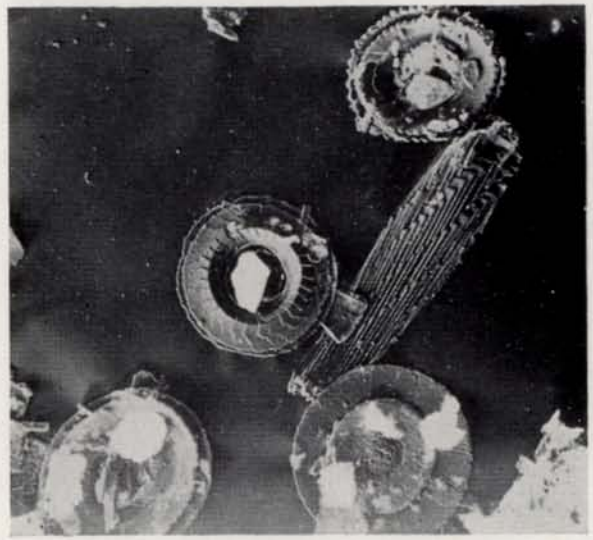


FIG. 10.14. *Replica of chalk particles.*

Comer, 1953), bacterial spores (Bradley and Williams, 1957), larger micro-organisms (Harris and Bradley, 1956), and even to yeast cells (Bradley, 1955). The extraction replicas now widely used in metallurgy (Fisher, 1954; Smith and Nutting, 1956) are essentially of the same type. The metal surface, coated with the replica material, is exposed to an etchant which selectively removes the matrix and leaves the particles of interest (precipitates, inclusions, etc.) adhering to the replica film (Figure 10.15).

Replicas of Biological Material

Most of the methods described above may also be used on biological specimens, provided that they have been

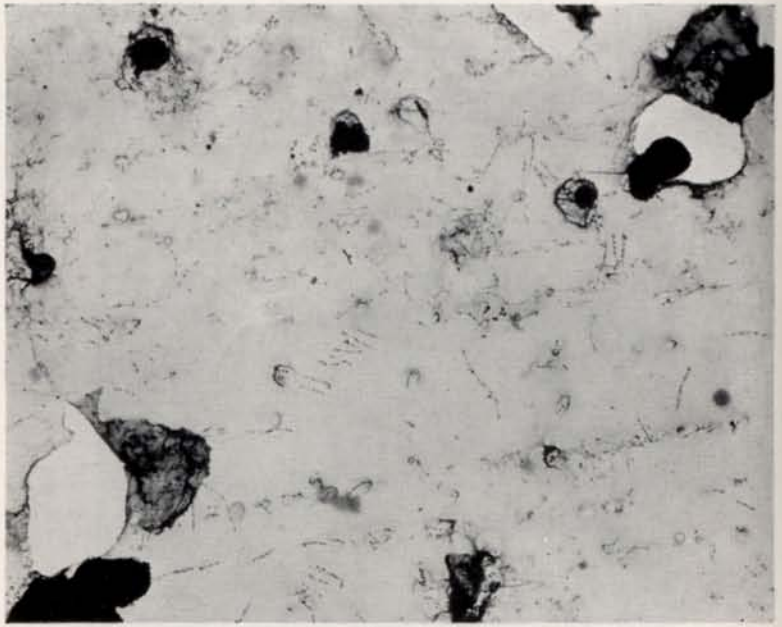


FIG. 10.15. *Extraction replica of high temperature steel, aged for 20 hours at 700°C, showing niobium carbide precipitates on dislocations. (M. J. Arrowsmith and Dr. J. Nutting, Department of Metallurgy, Cambridge University.)*

previously dried down, preferably after fixation. Details of a variety of applications have been given by Bradley (1956, 1958). Some workers prefer to put down a very thin layer of shadowing metal before evaporating the carbon on the specimen surface ('pre-shadowed replica'), in order to improve contrast. Labaw and Wyckoff (1955, 1958) have obtained replicas showing striking detail in plant virus crystals by this means (Figure 10.16). Various replica methods have been proposed for application to biological material in the wet state, but none of them has proved capable of producing reliable results except, perhaps, the three-stage method of Fischbein (1950). The first replica is made in a film of polyethylene, 0·1 mm thick, the second

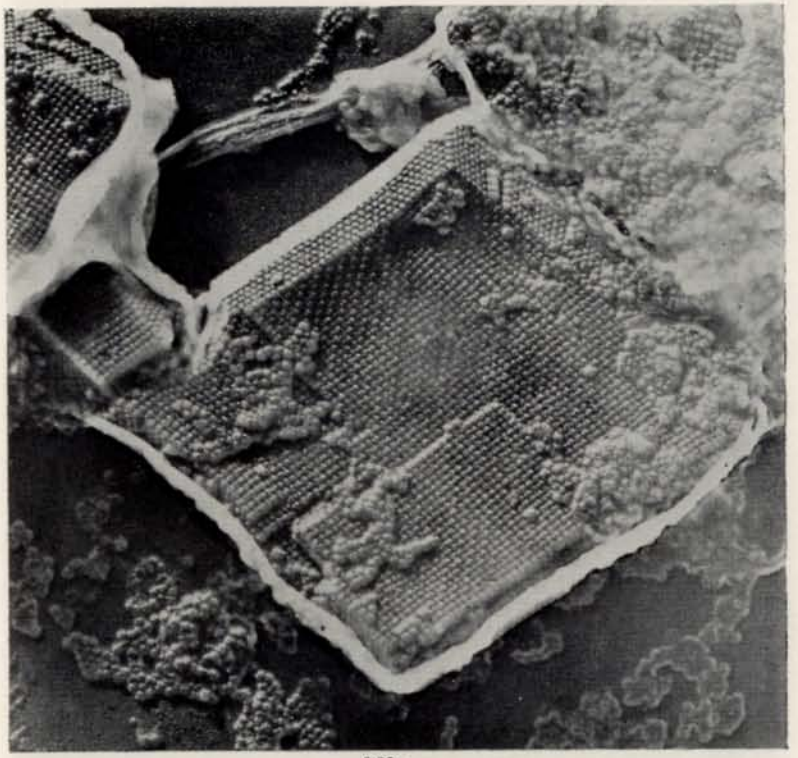


FIG. 10.16. *Carbon replica (with platinum stabilising layer) of tobacco necrosis virus crystal (L. W. Labaw and R. W. G. Wyckoff, National Institutes of Health, Bethesda).*

by painting this with a 1% solution of polystyrene in benzene, and the third by evaporating silica. Satisfactory replicas of undried surfaces of wood were obtained by this method.

For investigating at very high resolution structures which might be disturbed in the drying process, a replica process for frozen specimens was devised by Hall (1950). The wet specimen is rapidly cooled to liquid nitrogen temperature and then slowly raised to that of solid CO_2 (about $-75°C$), so that ice sublimes from the surface layers. After this 'etching' process has gone far enough

0·2μ

FIG. 10.17. *Low-temperature replica of polio virus crystal (R. L. Steere and F. L. Schaffer, Virus Laboratory, Berkeley, California).*

to reveal detail of the order suspected to exist, the surface is shadowed with chromium *in situ*, and this is followed by a supporting layer of silicon monoxide. After raising to room temperature and removing from the vacuum, the replica film is floated off in water and collected on a grid. An improved method, using carbon instead of silicon monoxide and followed by a backing film of plastic, has been described by Steere and Schaffer (1958), who obtained replicas of crystals of plant viruses and of the polio virus with it (Figure 10.17).

Biological Applications

Electron microscopical methods are now being applied to almost all types of biological problems in which the light microscope has shown the value of morphological studies. The main exception is the study of living material, which has so far proved impossible to study in the electron microscope. In most respects the methods of preparation of biological specimens (cf. T. F. Anderson, 1956; Sjöstrand, 1956) do not differ from those needed for normal light microscopy (in regard to fixation and dehydration) except that, finally, the section, smear or suspension has to be completely dried for insertion in the electron microscope. The possibility of artefacts being introduced at this stage has not been investigated as thoroughly as might be wished, but the evidence that has accumulated tends to show that the damage is not so serious as was initially feared. A considerable amount of work has been done by the freeze-drying technique, and by the related 'triple-point' method of T. F. Anderson (1956), which are generally agreed to be the best means of preserving biological structures during fixation and desiccation. No appreciable difference is usually to be seen between frozen-dried preparations and those fixed by other means, although opinions still differ on the merits of using osmium tetroxide, which was formerly much favoured for electron microscopy. Milder fixatives, such as potassium permanganate, are now increasingly used.

Detailed attention has also been paid recently to the development of staining methods adapted to the requirements of electron microscopy, for which the reagent must be of high electron scattering power, i.e. must contain a heavy element. Surveys of the problems involved have been made (Zeitler and Bahr, 1957, 1959; Cosslett, 1958; Valentine, 1958, 1959) and some work on positive staining methods has appeared (Sikorski and Simpson, 1958; Gersh, 1960; Barrnett, 1960). A negative staining procedure, in which the specimen is outlined by a dense material (phosphotungstic acid) which does not penetrate it, has been used by

Brenner and Horne (1959) to investigate the fine structure of bacteriophage (see Figure 10.21 later).

In almost all types of tissue and organism examined, the electron microscope has revealed a wealth of fine detail beyond the resolving power of the optical microscope, and largely unsuspected (cf. Wyckoff, 1958). Outside the cell nucleus a complex of lamellae, canaliculi, pores and membranes are found, frequently containing very small particles ('microsomes') (Figure 10.18). The larger inclusions, such as the mitochondria, already known from optical work, have been found to have an internal structure of similar complexity. A whole new field of micro-anatomy has been opened up which, at this level of size, cannot be separated from that of micro-physiology and biochemistry. The cell nucleus alone has so far resisted investigation, usually appearing almost uniformly dense, with denser formless regions here and there (cf. Figure 10.8). Experiments with different methods of fixation, however, have now led to the observation of chromosome structure in certain types of cell (cf. Pappas, 1956; Porter, 1960; Moses, 1960). Much attention is now being directed to finding the best means of preserving the nucleus for electron microscopy, now that the cytoplasm is well understood.

Most of the biological material examined with the electron microscope has been animal tissue, but botanical applications are increasing in number. Apart from the great amount of work done on fibre structure, noted above, the study of the plant cell wall (cf. Preston, 1958), of chloroplasts (Mühlethaler, 1955; McLean, 1960) and of leaf structure generally (Heitz, 1960) has received much attention. The leaf surface has also been recently explored by the carbon method (Juniper and Bradley, 1958; Juniper, 1960) (Figure 10.19).

Micro-organisms have also been shown to possess a quite complicated internal structure. Bacteria in particular have been much studied by the thin sectioning technique, aided by selective digestion and staining methods. Nucleic acid has been shown to be present in fairly well-defined regions, but it is not yet certain if bacteria possess a

FIG. 10.18. *Section of Pancreas, with endoplasmic reticulum, microsomes and* (below) *mitochondria* (*F. S. Sjöstrand, Karolinska Inst., Stockholm*).

FIG. 10.19. *Carbon replica of leaf surface* (*D. E. Bradley and B. E. Juniper*).

nucleus similar to that of higher organisms (Figure 10.20). The present state of knowledge has been reviewed by Kellenberger, Sechaud and Ryter (1960). These, and other, workers have also been studying the details of the infection of bacterial cells with bacteriophage and the subsequent development of a new generation of phage. By this means it has proved possible to reveal the unexpectedly complex structure of the phage particle itself (Brenner, 1960), shown in Figure 10.21. Since the different strains of phage offer one of the simplest systems for experiments in breeding, it is possible that evidence may be forthcoming of the

FIG. 10.20. *Section of Coli bacteria, embedded in polyester, showing internal structure* (*E. Kellenberger, Department of Biophysics, University of Geneva*).

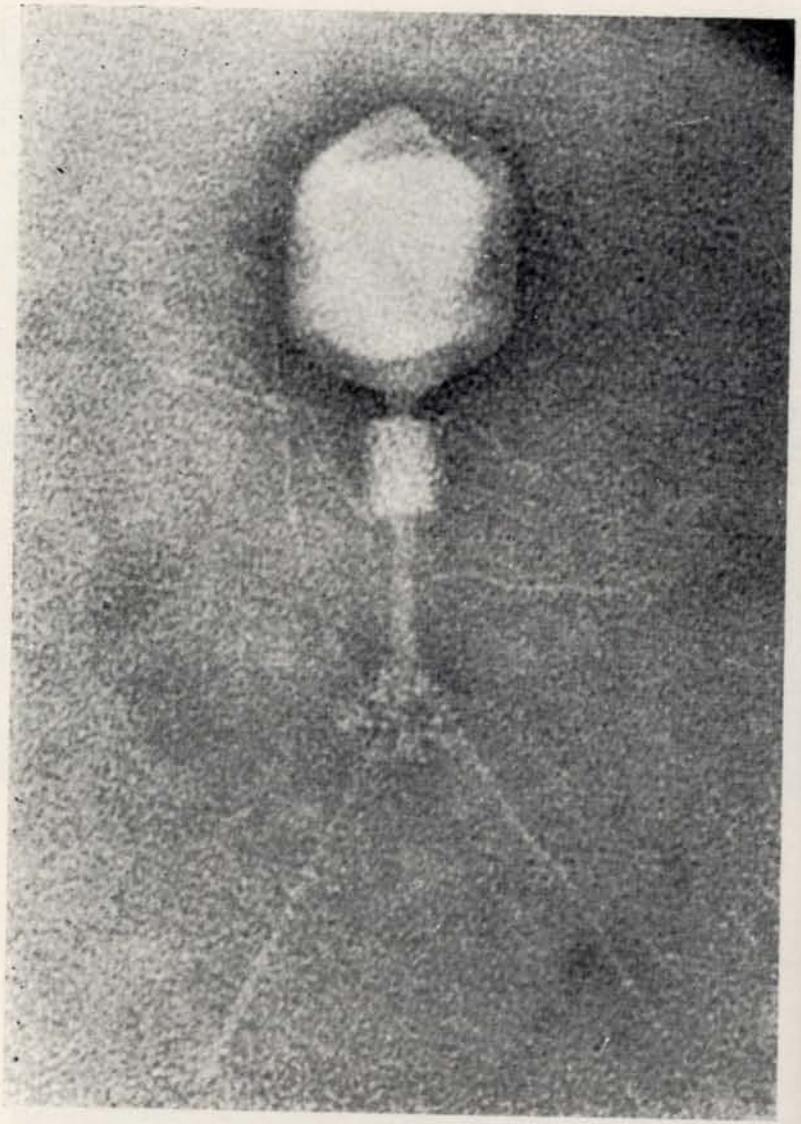


FIG. 10.21. *Bacteriophage of B. Coli, with phosphotungstic acid negative stain to show substructure (S. Brenner and R. W. Horne, Cavendish Laboratory).*

morphological basis of 'crossing-over' and other genetic effects.

The study of viruses is, if anything, even more indebted to electron microscopy for its advancement. All except the very largest, such as the influenza virus, lie beyond the resolution limit of the optical microscope. For the electron microscope they are 'big' particles, and a great deal of work has been done in measuring and classifying viruses

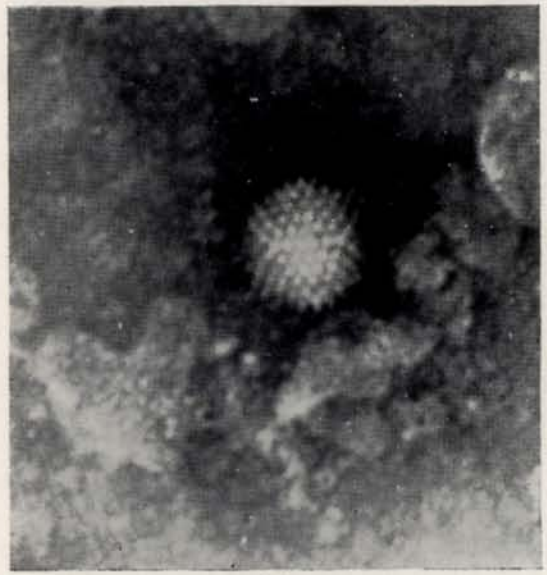


FIG. 10.22. *Adeno-virus, after treatment in phosphotungstic acid, showing sub-units arranged on the faces of an icosahedron (R. W. Horne, S. Brenner, A. P. Waterson and P. Wildy, Cavendish Laboratory and Pathology Department, Cambridge).*

according to their size and appearance. The larger viruses have been shown to have substructure (Figure 10.22) and even some of the smallest plant viruses are clearly not simple spheres or rods (Figure 10.23). Many plant viruses have been shown to form regular crystals, whose structure has been investigated by Wyckoff in particular (cf. Wyckoff,

FIG. 10.23. *Sugar beet virus, in phosphotungstic acid, showing spiral structure (R. W. Horne, A. R. Trim and G. Russell, Cavendish Laboratory and Plant Breeding Institute, Cambridge).*

1958). The polio virus has now also been found to crystallise (cf. Figure 10.17), and the electron microscope has played an important part in the isolation and identification of this and other pathogenic viruses.

So far the electron microscope has been utilised mainly in the fundamental aspects of biological research, but it is now beginning to find a place in the more practical branches of medicine and dentistry. There can be no doubt that the great amount of work going on in the investigation of the ultra-structure of all types of cells and tissues will lead to a much deeper understanding of the macroscopic features of health and disease. Nevertheless, it is likely to be a long time before the electron microscope becomes the regular laboratory tool which the optical microscope has long been in hospital as well as research laboratory. It may require even longer to overcome the very difficult problems connected with the observation of living specimens, which seems to be inherently opposed by a sort of uncertainty principle of biological instead of physical nature. The life process is so sensitive to desiccation, heat and electron bombardment, all of which are unavoidable in electron microscopes as now known, that the very attempt to make an observation must fatally distort the organism. Even if this fundamental obstacle is never to be overcome, however, electron microscopy has an almost unlimited field of application before it in the study of isolated biological material.

Metallurgical Applications

Until recently the use of the electron microscope in metallurgy was almost entirely limited to surface studies, by means of the replica process. The early work has been surveyed in a Symposium on Metallurgical Applications of the Electron Microscope (1950), and in the reports of the American Society for Testing Materials, quoted on p. 252. The later development of the extraction replica technique (see p. 255), by which certain constituents of the sample can be removed from the metal and examined in the electron microscope, greatly extended applications in metallurgy. By these methods a wide variety of processes has been investigated: precipitation, segregation, diffusion, sub-grain structure, fracture, surface wear, plastic deformation and so on. Miniature furnaces for following the

effect of high temperature, including quenching, have been described by Jones (1954), Itoh, Itoh and Watanabe (1954), Whelan (1960) and others. A straining stage has been made by Weichan (1955) and by Wilsdorf (1958).

Although the scope of electron microscopy was further widened by the introduction of the reflexion and scanning methods (see Chapter IX), all such surface studies remained open to the objection that they were not necessarily typical of the bulk structure of the metal. Techniques have now become available, however, for preparing most metals in the form of films thin enough for direct examination in the electron microscope, starting from bulk material (see p. 238). Some applications to light alloy problems, particularly the formation of Guinier-Preston zones, have been described by Nicholson, Thomas and Nutting (1958, 1959). By this technique almost all the standard procedures of metallurgy are now open to investigation at the highest resolution. The formation and movement of dislocations (Figure 10.7; Figure 10.24), the precipitation and resolution of constituents, and other processes have been followed continuously and even filmed from the screen of the electron microscope. Up-to-date surveys of progress have been made by Hirsch (1959) and Franks and Revell (1959); see also the group of papers in the August 1959 issue of *J. Inst. Metals.*

Crystals and Crystal Lattices

The study of the structure, and especially the growth, of crystals was initially also limited to the use of replica techniques. Later, both Hall (1950a) and Dawson (1952) developed methods of growing micro-crystals of organic substances, such as long-chain paraffins, proteins and other macro-molecules, directly on the supporting grid. In this way Dawson was able to investigate the occurrence of spiral dislocations and other types of growth processes. Hall resolved the lattice structure of a number of proteins, down to a spacing of about 50 Å, that is, on a much finer scale than the virus crystals studied by Wyckoff and shown in Figure 10.17. Subsequent work by Menter (1956) with a high resolution microscope succeeded in resolving the

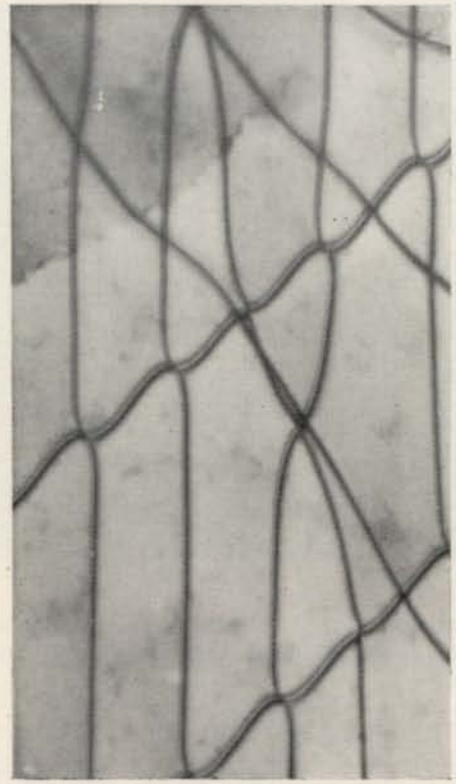


FIG. 10.24. *Parallel arrays of dislocations in bismuth telluride film (G. A. Geach and R. Phillips, A.E.I. Research Laboratory, Aldermaston).*

lattices of copper and platinum phthalocyanine, with spacings down to 10 Å. More recently Bassett and Menter (1957) have imaged the 020 spacing of molybdenum trioxide (6·9 Å) (Figure 10.25). The image is formed by diffraction from planes which lie almost parallel to the direction of incidence. When several sets of planes give diffracted beams, great care is necessary in interpreting the image, since spurious periodicities may appear (Kamiya, Nonoyama, Tochigi and Uyeda, 1960).

When two lattices are superimposed, Moiré effects are

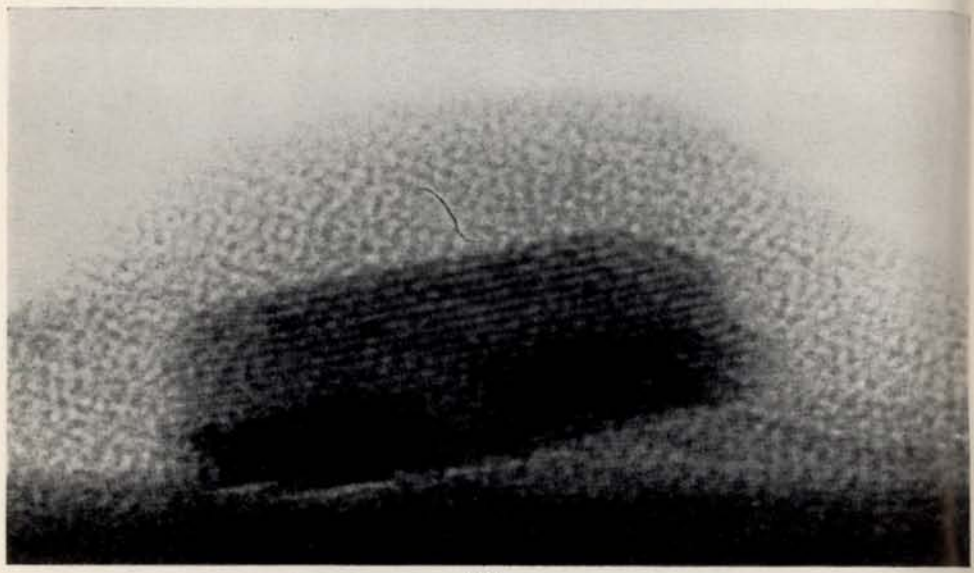


FIG. 10.25. *Direct resolution of 020 spacing in molybdenum trioxide lattice (G. A. Bassett and J. W. Menter, Tube Investments Research Laboratory, Hinxton).*

observed, as first noted by Mitsuishi, Nagasaki and Uyeda (1951); they obtained fringes about 100 Å apart from lamellar crystals of graphite. The detailed explanation in terms of crossed lattices was given by Dowell, Farrant and Rees (1954, 1956). Moiré patterns may also arise from the existence of two lattices of slightly different spacing within the same crystal, if two phases co-exist or diffusion is occurring. Alternatively, a metal may be deposited by evaporation on to a film of another metal, when fringes will appear which are characteristic of the difference in their respective lattice constants and the angle which one makes with the other; Figure 10.26 shows a typical result with palladium on gold (Bassett, Menter and Pashley, 1958). The experimental techniques and theoretical treatment of the imaging of crystal lattices have been surveyed by Menter (1958). Several interesting new applications of the

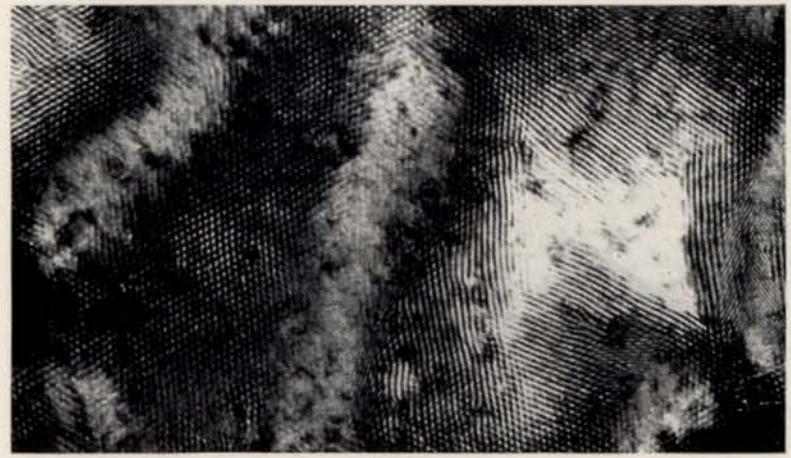


FIG. 10.26. *Moiré pattern in composite film of palladium-gold (G. A. Bassett, J. W. Menter and D. W. Pashley, Tube Investments Research Laboratory, Hinxton).*

electron microscope have been opened up by this work, among them the investigation of super-lattice structures (Glossop and Pashley, 1959).

REFERENCES

AGAR, A. W. *Brit. J. Appl. Phys.*, **7** (1956), 17.
ANDERSON, N. G. *Physical Techniques in Biological Research*, Academic Press, **III**, Chap. 7 (New York, 1956).
ANDERSON, T. F. *Physical Techniques in Biological Research*, Academic Press, **III**, Chap. 5 (New York, 1956).
BACHMANN, L. and SITTE, P. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 75.
BACKUS, R. C. and WILLIAMS, R. C. *J. Appl. Phys.*, **21** (1950), 11.
BARRNETT, R. J. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 91.
BASSETT, G. A., MENTER, J. W. and PASHLEY, D. W. *Proc. Roy. Soc.*, A.**246** (1958), 345.
BASSETT, G. A. and MENTER, J. W. *Phil. Mag.*, **2** (1957), 1482.
BRADLEY, D. E. *Brit. J. Appl. Phys.*, **5** (1954), 65 and 96.
BRADLEY, D. E. *Research Correspondence*, **8** (1955), S.62.

BRADLEY, D. E. *J. Appl. Phys.*, **27** (1956), 1399.
BRADLEY, D. E. *Nature*, **181** (1958), 875.
BRADLEY, D. E. *J. Roy. Micr. Soc.* **79** (1959/60), 101.
BRADLEY, D. E. *Brit. J. Appl. Phys.*, **10** (1959), 198.
BRADLEY, D. E. and WILLIAMS, D. J. *J. Gen. Microbiol.*, **17** (1957), 75.
BRANDON, D. G. and NUTTING, J. *Brit. J. Appl. Phys.*, **10** (1959), 255.
BRENNER, S. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 621.
BRENNER, S. and HORNE, R. W. *Biochim. Biophys. Acta*, **34** (1959), 103.
CALBICK, C. J. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **I**, 422.
CASTAING, R. *Revue Univ. Mines*, **12** (1956), 405.
COSSLETT, A. and COSSLETT, V. E. *Brit. J. Appl. Phys.*, **8** (1957), 374.
COSSLETT, A. and COSSLETT, V. E. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. I, 518.
COSSLETT, V. E. *J. Roy. Micr. Soc.*, 78 (1958), and 18.
COSSLETT, V. E. and HORNE, R. W. *Vacuum*, **5** (1957), 109.
DAWSON, I. M. *Proc. Roy. Soc.*, A.**214** (1952), 72.
DOWELL, W. C., FARRANT, J. L. and REES, A. L. G. *Proc. Third Internat. Conf. on Electron Microscopy, London* (1954), Roy. Micr. Soc., London, p. 279; *Proc. Tokyo Conf. on Electron Microscopy* (1956), Japanese Society for Electron Microscopy, p. 320.
DRUMMOND, D. G. *J. Roy. Micr. Soc.*, **70**, Part 1 (1950), 1–141.
FERT, C., MARTY, B. and SAPORTE, R. *C.R. Acad. Sci.*, Paris, **240** (1955), 1975.
FISCHBEIN, I. W. *J. Appl. Phys.*, **21** (1950), 1199.
FISHER, R. M. *A.S.T.M. Special Technical Publication*, No. 155, (1954).
FRANKS, A. and REVELL, R. S. M. *Brit. J. Appl. Phys.*, **10** (1959), 438.
GEACH, G. A. and PHILLIPS, R. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **I**, 571.
GERSH, I. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 89.

GLAUERT, A. M. and GLAUERT, R. H. *J. Biophys. Biochem. Cytol.*, **4** (1958), 191.
GLAUERT, A. M., ROGERS, G. E. and GLAUERT, R. H. *Nature*, **178** (1956), 803.
GLOSSOP, ANN B. and PASHLEY, D. W. *Proc. Roy. Soc.*, A.**250** (1959), 132.
GREEN, P. and CHAPMAN, G. *Amer. J. Bot.*, **42** (1955), 685.
HALL, C. E. *J. Appl. Phys.*, **21** (1950), 61.
HALL, C. E. *J. Biol. Chem.*, **185** (1950a), 45 and 749.
HAMM, F. A. and COMER, J. J. *J. Appl. Phys.*, **24** (1953), 1495.
HARRIS, K. and BRADLEY, D. E. *J. Roy. Micr. Soc.*, **76** (1956), 37.
HEIDENREICH, R. D. and PECK, V. G. *J. Appl. Phys.*, **14** (1943), 23.
HEIDENREICH, R. D. *J. Appl. Phys.*, **20** (1949), 993.
HEITZ, E. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 499 and 501.
HIRSCH, P. B. *Metallurgical Reviews*, **4** (1959), 101.
HIRSCH, P. B., HORNE, R. W. and WHELAN, M. J. *Phil. Mag.*, **1** (1956), 677.
HOLLAND, L. *Vacuum Deposition of Thin Films*, Chapman and Hall (London, 1956).
ITOH, K., ITOH, T. and WATANABE, M. *Proc. Third Internat. Conf. on Electron Microscopy, London* (1954), Roy. Micr. Soc., London, 658.
JONES, D. *Proc. Third Internat. Conf. on Electron Microscopy, London* (1954), Roy. Micr. Soc., London, 662.
JUNIPER, B. E. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 489.
JUNIPER, B. E. and BRADLEY, D. E. *J. Ultrastructure Res.*, **2** (1958), 16.
KAMIYA, Y., NONOYAMA, M., TOCHIGI, H. and UYEDA, R. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **I**, 339.
KELLENBERGER, E., SECHAUD, J. and RYTER, A. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 212.
KELLY, P. M. and NUTTING, J. *J. Inst. Metals*, **87** (1959), 385.
LABAW, L. W. and WYCKOFF, R. W. G. *Exptl. Cell Res.*, **3**, Suppl. (1955), 395.
LABAW, L. W. and WYCKOFF, R. W. G. *J. Ultrastructure Res.*, **2** (1958), 8.
LACY, D. (Ed.). *The Microtomist's Vade Mecum*, Churchill (London, 1960).

MENTER, J. W. *Proc. Roy. Soc.*, A.**236** (1956), 119.
MENTER, J. W. *Phil. Mag. Suppl.*, **7**, No. 27 (1958), 299.
Metallurgical Applications of the Electron Microscope, Institute of Metals (London, 1950).
MCLEAN, J. D. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 27.
MITSUISHI, T., NAGASAKI, H. and UYEDA, R. *Proc. Imp. Acad., Japan*, **27** (1951), 86.
MOSES, M. J. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 199 and 230.
MÜHLETHALER, K. *Internat. Rev. Cytology*, **4** (1955), 197.
NICHOLSON, R. B., THOMAS, G. and NUTTING, J. *Brit. J. Appl. Phys.*, **9** (1958), 25; and *J. Inst. Metals*, **87** (1959), 429.
NIXON, H. L. and FISHER, H. L. *Brit. J. Appl. Phys.*, **9** (1958), 68.
PAPPAS, G. D. *J. Biophys. Biochem. Cytol.*, **2** Suppl. (1956), 431.
PEACHEY, L. D. *J. Biophys. Biochem. Cytol.*, **4** (1958), 233.
PEACHEY, L. D. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 72.
PORTER, K. R. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 186.
PRESTON, R. D. *Nature*, **182** (1958), 1402.
REIMER, L. *Elektronenmikroskopische Untersuchungs- und Präparationsmethoden*. Springer, Berlin (1959).
ROCHOW, T. G. and BOTTY, M. C. *Analyt. Chem.*, **30** (1958), 640.
ROCHOW, T. G., THOMAS, A. M. and BOTTY, M. C. *Analyt. Chem.*, **32** (1960), No. 5, 92R.
RYTER, A. and KELLENBERGER, E. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **II**, 52.
SCHLIPKÖTER, H. W., STEIGER, H., ESSER, H. F. and BECK, E. G. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **I**, 754.
SIKORSKI, J., *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **I**, 686.
SIKORSKI, J. and SIMPSON, W. S. *Nature*, **182** (1958), 1235.
SITTE, H. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1959), Vol. **II**, 63.
SJÖSTRAND, F. S. *Physical Techniques in Biological Research*. Academic Press, **III**, Chap. 6 (New York, 1956).
SMITH, E. and NUTTING, J. *Brit. J. Appl. Phys.*, **7** (1956), 214.
STEERE, R. L. and SCHAFFER, F. L. *Biochim. Biophys. Acta.*, **28** (1958), 241.

SWERDLOW, M. *Analyt. Chem.*, **26** (1954), 34.
SWERDLOW, M., DALTON, A. J. and BIRKS, L. S. *Analyt. Chem.*, **28** (1956), 597.
TOLANSKY, S. *Multiple Beam Interferometry* (Oxford, 1948).
TURIAN, G. and KELLENBERGER, E. *Exptl. Cell Res.*, **11** (1956), 417.
VALENTINE, R. C. *J. Roy. Micr. Soc.*, **78** (1958), 26; *J. Pho. Sci.*, **7** (1959), 66.
WEICHAN, C. *Z. Wiss. Mikroskop.*, **62** (1955), 147.
WHELAN, M. W. *Proc. Fourth Internat. Conf. on Electron Microscopy, Berlin, 1958*. Springer, Berlin (1960), Vol. **I**, 96.
WILSDORF, H. G. F. *Rev. Sci. Instrum.*, **29** (1958), 323
WYCKOFF, R. W. G. *The World of the Electron Microscope*. Yale Univ. Press (New Haven, 1958).
ZEITLER, E. and BAHR, G. F. *Exptl. Cell Res.*, **12** (1957), 44; *J. Appl. Phys.*, **30** (1959), 940.

Subject Index

Author Index

HAINE

ELECTRON MICROSCOPE

SPON